AN EXPLORATION OF ALTERNATIVE GULF OF ALASKA DOVER SOLE ASSESSMENT MODELS

Carey McGilliard, August 2013

INTRODUCTION

The purpose of this document is to outline a proposed change from conducting assessments using the previously used Dover sole assessment model framework to conducting assessments using Stock Synthesis version 3.24o (SS3; Methot and Wetzel 2013).

Previous assessments were conducted using an ADMB-based age- and sex-structured population dynamics model with length-at-age, weight-at-length, maturity-at-age, and age-length transition matrices estimated outside of the model. The previous model estimated the log of mean recruitment, parameters for logistic age- and sex-specific selectivity curves for the fishery and survey, recruitment deviations, and yearly fishing mortality rates. The model included ages 3-40 (age 40 was a plus group) and excluded data for fish below age 3 and 18cm in length.

SS3 is a flexible assessment model framework that extends the capabilities of the 2011 Dover sole assessment model to address the concerns of the GOA Plan Team, the SSC, and previous Dover sole assessment authors. Although we do not expect that all concerns can be addressed within the time-frame for the 2013 assessment cycle, this document outlines the work that was done to transition the Dover sole assessment from the previous assessment framework to SS3. In addition, proposed alternative models that address some previous concerns about the Dover sole assessment by using the extensive suite of modeling options available in SS3 are discussed.

SSC AND PLAN TEAM COMMENTS ON PREVIOUS ASSESSMENTS

In 2011, Gulf of Alaska (GOA) Dover sole was managed as a Tier 5 species on the recommendation of the assessment authors due to decreased confidence in the 2011 and 2009 age-structured assessment models. MCMC analysis conducted in 2011 showed that the likelihood for the accepted 2009 model was a local maximum (Stockhausen et al. 2011).

Previous assessment authors suggested that growth rates, natural mortality rates, and age and size classes used in the model be re-evaluated. In addition, authors suggest that alternative selectivity functions be explored and that ageing error and internal estimation of growth be considered.

Two currently unfulfilled SSC requests exist:

1. SSC comment: "*Because adjacent age-classes are likely to overlap in size and spatial distribution, the fishery selectivity curves estimated by the model seem implausibly steep, possibly indicating mis-specification of the age-length conversion matrices. The SSC requests that the growth model and age-length conversion matrices be re-evaluated in the next assessment.*"
2. SSC request: *The SSC requested that the next round of assessments consider the possible use of ADF&G bottom trawl survey data to expand the spatial and depth coverage.*

The previous framework for conducting Dover sole assessments was unable to address these concerns, but these can be readily explored using SS3. Relative to the 2011 model, SS3 offers the following features:

(1) The 2011 assessment found that the 2009 assessment had reached a “local minimum” for the objective function. SS3 offers a “jitter” option, which allows for initial parameter values to be adjusted by a random deviate. Iteratively running the model with the “jitter” option turned on allows the user to start the model from a wide range of initial values so as to identify the best objective function value.
(2) A request concerning the previous Dover sole assessments was that the age-length transition matrices and other growth parameters be re-examined and potentially estimated within the model. The 2011 model had limited capability to do this but such flexibility is included in the SS3 framework.
(3) Mean weight-at-age data can be included in the SS3 model and can be used as a likelihood component to help estimate growth. Since these data are available for GOA Dover sole their use within the assessment model would be advantageous.
(4) SS3 has many options for specifying the functional form of selectivity curves and these could be used to explore length-based fishery selectivity for Dover sole, which may be a more accurate reflection of the selection process than the knife-edge, age-based fishery selectivity estimated in previous assessments.
(5) SS3 allows for specification of ageing error. Ageing error is ignored in the current model, but Dover sole are known to be one of the harder species to age (Abookire and Macewicz 2003).
(6) SS3 allows for multiple survey and fishing fleets to be included in the model. This feature would be needed to explore the inclusion of the ADF&G bottom trawl survey in future assessments; the previous model accommodated only one fishery and one survey.
(7) SS3 accommodates age-composition data for ages 0-2. The previous assessment model omitted data for fish below age 3. Including data for ages 0-2 may inform recruitment estimates and age-based selectivity at young ages.
(8) SS3 allows for calculation of mid-year weight-at-age which is an improvement over the 2011 model because it more accurately matches biological processes that occur during the year with timing of fishing.
(9) The previous assessment model assumed the stock was unfished prior to the model start year, but we know that fishing occurred before 1984. SS3 allows the user to estimate an initial fishing mortality rate to account for fishing prior to the availability of catch data.
(10) SS3 is used by many scientists worldwide, which provides an ad-hoc quality control system for identifying bugs in the code.

ANALYTIC APPROACH: TRANSITION OF 2011 MODEL INTO AN EQUIVALENT SS3 MODEL

Matching population dynamics between models

Mean recruitment

Several steps were taken to build an SS3 model with population dynamics that matched those of the 2011 model using deterministic models with no estimation of parameters and no recruitment deviations. First, the relationship between the log of mean recruitment estimated in the 2011 model $(\ln(\bar{R}))$ and the log of R_0 (unfished recruitment $(\ln(R_0))$ that is estimated in SS3 was determined (Equation 1), where M is natural mortality.

(1) $$\ln(R_0) = \ln\left(\frac{2\bar{R}}{1000}\right) + 3M$$

The $\ln(\bar{R})$ estimated in the 2011 model refers to female mean recruitment of age 3 individuals, while $\ln(R_0)$ refers to total recruitment (males and females) of age 0 individuals in thousands; both models assume a 1:1 sex ratio (but any sex ratio can be specified in SS3; a different sex ratio would change Equation 1). Using Equation 1, equivalent deterministic runs were conducted with fixed parameters at their maximum likelihood estimates (MLEs) from the 2011 model. This was to ensure that both models had the same behavior in the absence of estimation. Equation 1 was required to ensure that numbers at age 3 and above are the same in both models for an unfished population.

Selectivity

The 2011 model assumed sex-specific age-based logistic selectivity functions for fishery and survey selectivity. Although SS3 has logistic, sex-specific selectivity, it was found that the specification of male logistic age-based selectivity in SS3 was difficult to cast into a logistic shape. Sex-specific length-based logistic selectivity can be specified such that selectivity can be estimated for both sexes while retaining the logistic shape, or age-based double normal selectivity curves could be specified with a large value for the standard deviation of the descending limb such that asymptotic, logistic-like, sex-specific selectivity could be estimated. In the interest of matching the 2011 model as closely as possible, the age-based, sex-specific double normal selectivity curves with no descending limbs were used for fishery and survey selectivity curves. The fishery selectivity curves in SS3 were matched as closely as possible to the age-based logistic curves from the 2011 model for the purpose of comparing population dynamics between the models and are a near-exact match (but were logistic for the 2011 model and double-normal for the SS3 model; Figure 1). Deterministic runs conducted for Dover sole using the fishery selectivity curves in Figure 1 led to the same time series of SSB for both models (Figure 2), indicating that the population dynamics of the models are the same. Figure 4 shows an example of double-normal selectivity curves that match the shape of the logistic curves from the 2011 model to some degree. The slight mis-specification of selectivity curves in SS3 results in small differences in population dynamics between the 2011 and SS3 models that are evident in the estimates of SSB over time (Figure 5).

Stock-Recruitment

The 2011 model estimated recruits as median-unbiased recruitment deviations from their mean value. The SS3 model was configured similarly by specifying a Beverton-Holt stock-recruitment curve with a steepness of 1. SS3 estimates mean-unbiased recruitment deviations by specifying σ_R and applying a bias adjustment factor. For the deterministic runs, σ_R was set to 1.0E-06, and for runs when recruitment deviations were estimated, σ_R was set to 0.49. The 2011 model estimated recruits (age 3) freely (i.e. no σ_R) and this constitutes a difference between the models.

Growth

The 2011 model used empirical estimates of maturity-at-age sex-specific somatic weight-at-age. SS3 also can use similar empirically specified values for the calculation of spawning stock biomass and biomass-at-age (Figure 6). A benefit of using the SS3 framework is the ability to specify and estimate growth parameters internally. When growth parameters are specified (instead of age-specific schedules), small differences arise between models because SS3 uses the beginning of the year weight-at-age to calculate SSB (like in the 2011 model), but uses mid-year weight-at-age to calculate exploitable and survey biomass (the 2011 model uses beginning-of-the-year weight-at-age for all calculations).

In addition, age-length transition matrices were specified directly in the 2011 model whereas in SS3 they are computed from specified von-Bertalanffy growth curve parameters and CVs in length-at-age. To match population dynamics between models, the CVs of the youngest and oldest age classes were estimated externally and specified within SS3. The resulting age-length transition matrices output from SS3 runs were examined to check that they closely matched those used in 2011. A request concerning the previous Dover sole assessments was that the age-length transition matrices and other growth parameters be re-examined and potentially estimated within the model. SS3 provides ample flexibility to explore growth relationships whereas this option was unavailable in the 2011 model.

Biomass

Differences in total biomass will occur between the models because SS3 includes ages 0-2. However, SSB and survey biomass were shown to be matched precisely between models when run deterministically when selectivity curves match between models and other parameters are fixed (Figure 2 and Figure 3).

Timing

Both the SS3 and 2011 model calculated spawning stock biomass, survey biomass, and recruitment at the beginning of the year. SS3 calculates exploitable biomass in the middle of the year, but a vector for weight-at-age was manually provided to SS3 which forced the model to use beginning-of-the year weight-at-age in the exploitable biomass calculation to match the 2011 model as closely as possible.

Data used in SS3 and the 2011 Model

The same data used in the 2011 Dover sole assessment model (Stockhausen et al. 2011, page 758) were used in the SS3 model: survey biomass, survey age- and length-compositions (triennial for 1984-1999 and biennial for 2001-2011), fishery length-composition data (1985-2011), and catch history (1984-2011). An important difference between the 2011 model and SS3 is that the youngest age class in the 2011 model (age 3) represents only age 3 individuals, while SS3 population dynamics begin at age 0 and consider the lowest age and length bins of data to be the proportion of individuals ages 0-3 and lengths 0-the upper limit of the lowest length bin, respectively. Therefore, age- and length-composition data must include ages 0-2 and any lengths no matter how small in SS3, while the 2011 model omitted data on ages 0-2 (and excluded data on fish smaller than 18cm). That SS3 included data on ages 0-2 likely informs estimates of selectivity at the lowest ages and hence improves recruitment estimates (especially in the most recent years). Ignoring this difference between models will result in extreme differences between expected and observed age- and length-compositions for the youngest age and length bins when selectivity at these ages and lengths is greater than 0. An alternative solution to including additional data in SS3 model runs was to specify an additional selectivity-at-length curve as a knife-edge curve with selectivity equal to zero at lengths where fish are likely to be younger than age 3 (in SS3 it is possible to specify selectivity-at-age and at-length at the same time). This was a coarse solution, as fish at age 3 are a variety of lengths and it required internal specification of growth parameters, which meant that maturity-at-age and weight-at-age would not be an exact match between the 2011 model and the SS3 model. Therefore, the SS3 model was set up to match the 2011 model, but included data on proportions at ages 0-2. Likewise, proportions at lengths smaller than 18cm were included in the lowest (18-20cm) length bin.

In 1990, 1993, 1996, and 2001, surveys covered a more restricted depth range than in other years. This was handled in the 2011 model by inflating survey biomass estimates by year-specific availability factors in years when only shallower water was surveyed and estimating a separate survey selectivity curve for those years. Likewise, in SS3, separate a separate selectivity curve was specified for the years when only shallower waters were covered and the same availability factors were used. This was accomplished by defining a second survey for those years.

Parameter Estimation in SS3 and the 2011 Model

Parameters Estimated Inside the Assessment Model

SS3 and 2011 model runs were conducted with estimation of the log of mean recruitment, recruitment deviations, fishing mortality rates (using the same empirical growth vectors in both models), and selectivity parameters. Selectivity parameters for the fishery, full coverage survey years, and shallow water survey years were estimated; the location of peak selectivity and the width of the ascending limb of the selectivity curve were estimated in SS3 and the age at 50% selection as well as the slope of the logistic selectivity curves were estimated in the 2011 model.

Likelihood component for survey biomass index

Table 2 lists the likelihood components used in SS3 and the 2011 model. The likelihood component for the survey biomass index and the data used to calculate the survey biomass likelihood component are the same for both models. The 2011 model and SS3 survey biomass values match almost exactly in a deterministic model with no estimation (Figure 3).

Age- and length-composition likelihood components

The age- and length-composition likelihood components in SS3 are identical to those in the 2011 model. However, as noted above, the observations of survey proportions-at-age and proportions-at-length differ among models in that the data given to SS3 includes the data given to the 2011 model in addition to the proportions of age 0-2 fish and lengths below 18cm. Therefore, the values of these likelihood components cannot be compared directly between the 2011 model and SS3, but are expected to have similar influences on model fits. The fits to age- and length-composition data are very similar among models (Figure 13-Figure 15). The addition of age 0-2 and small length data included in the SS3 model likely contribute to differences in numbers at age 3 and selectivity parameter estimates. There is no easy way to test the extent to which the additional data contributes to differences, as the 2011 model does not accept the additional data, while it is required for the SS3 model.

Recruitment likelihood components

Recruitment likelihood components differ slightly between models. The 2011 model does not include a CV for recruitment deviations. Both models allow for estimating early-period (1947-1983), main-period (1984-2008), and late-period (2008-2011) recruitment deviations as separate likelihood components, but the 2011 model also includes the early period recruitment deviations in the likelihood component for the main-period (Table 2). There is no way to include early period recruits in both an early-period and main-period likelihood component in SS3. In the 2011 model, the recruitment deviations for the main and late time periods must sum to 0. The purpose of defining recruitment periods is that the recruitment deviations from one time period cannot influence the recruitment deviations from another time period by way of forcing the deviations to sum to 0. In SS3, only the main-period recruitment deviations have a formal sum-to-0 constraint, but it is expected that the early- and late-period recruits will come close to summing to 0. The likelihood components for early recruits have a weighting of 2x the value of that likelihood component and the late-period recruits have a weighting of 3x the value of the early-period likelihood component in the 2011 model, while both have a weighting of 1 in SS3. SS3 does not allow the user to adjust the weighting of the likelihood components for early-period recruitment deviations. Runs of the 2011 model with re-weighting of the early- and late-period recruitment likelihood components to 1 show that the likelihood weightings do not make a noticeable difference in estimation and model fits. The inclusion of early-period recruitment deviations as a separate likelihood component as well as part of the main-period recruitment deviations likely contributes to differences in initial numbers of recruits and SSB. Differences between models are smallest when including early-period recruits as a separate likelihood component and not in the main-period likelihood component in SS3, rather than vice versa. In addition, including early-period recruits as a separate likelihood component prevents the early-period recruitment deviations from influencing the values of main-period recruitment deviations; this is sensible because any fishing prior to 1984 is taken into account using early-period recruitment deviations (as the models assume that no fishing occurred prior to 1984) and thus tend to be negative and fewer data exist to inform early-period recruitment deviations.

ANALYTIC APPROACH: PROPOSED ALTERNATIVE SS3 MODELS

The following models are proposed alternatives to the transitional SS3 model that was constructed to match the dynamics of the 2011 model:

M0: The transitional SS3 model described above (the SS3 model that best matches the dynamics of the 2011 model)

M1: Length-based fishery selectivity. The fishery data consist only of length compositions and therefore the model may be able to estimate length-based selectivity more effectively than age-based selectivity. Fishing selectivity may be more a process of length (e.g. due to the net's mesh size) than age (where multiple ages of fish are the same length). SS3 is able to estimate length-based sex-specific logistic fishery selectivity, so there is no need to use a double-normal curve with no descending limb for this alternative.

M2: Estimate an initial equilibrium fishing mortality rate. The transitional SS3 model assumes that the stock was unfished prior to the model start year (1984) even though fishing occurred before 1984. In the transitional model, estimates of recruitment for years prior to 1984 were below average, which may be an artifact to account for fishing that occurred prior to 1984.

M3: Internal specification of growth parameters. The transitional SS3 model used empirical estimates of age-specific maturity and body weight. This model also was configured to have the same values to use at both the beginning *and* middle of the year. Internally specifying growth parameters allows the model to account for fish growth throughout the year by calculating weight-at-age in the middle of the year, which is used to calculate exploitable biomass.

M4: A combination of M1, M2, and M3, where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery selectivity is a logistic, sex-specific, length-based function.

M5: As for M4, but with length-based, logistic, sex-specific selectivity for the two surveys (as well as for the fishery).

Further proposed alternative models

The SS3 model framework facilitates the potential for the following analyses to be conducted:

- Adding mean weight-at-age data to the assessment and estimating growth parameters internally, given that there is a mismatch between the Abookire & Macewicz (2003) growth relationships and those used in the assessment model. Estimating growth in addition to parameters that are currently estimated in the transitional SS3 model without the addition of mean weight-at-age data resulted in poor fits to the data.
- Estimating growth parameters and the age-length transition matrix outside of the model, given the mismatch between the Abookire & Macewicz (2003) maturity ogive and von Bertlanffy growth curve and those used in the assessment model. Fitting the transitional SS3 model to the data using the Abookire & Macewicz (2003) growth relationships (including their weight-length relationship, which is already used in the transitional SS3 model) resulted in very a poor fit to the data.
- Including ageing error in the model: the previous assessment models ignored ageing error. The CVs about the length-at-age relationship are quite large. This implies that there are some age 3 fish that are the same length as some age 20+ fish, which is likely untrue and could potentially be attributed to ageing error.
- Re-evaluating effective sample sizes for age- and length-composition data. There are abrupt year-to-year changes in age-compositions that occur in the observations that are likely due to observation error. Using such high effective sample sizes may exclude some process errors which should be considered.
- Exploring alternative methods for handling years where the survey sampled only shallow water. The current method assumes that if more area were surveyed, the same biomass of fish per area would have been caught and the same proportions of ages and lengths would have been sampled. However, Dover sole moves ontogenetically and spatial dynamics are sex-specific. The shallow-water survey years are handled as a separate survey in SS3. Alternative models could explore estimating catchability or allowing for the estimation of dome-shaped selectivity for the shallow-water survey instead of adjusting survey biomass data points by an availability factor.

RESULTS: TRANSITION OF 2011 MODEL INTO AN EQUIVALENT SS3 MODEL

The 2011 and SS3 models each estimated a similar time series of numbers at age 3 (considered recruits in the 2011 model), but the SS3 model estimated fewer numbers at age 3 than the 2011 model starting in the late 1990s (Figure 7). Numbers at age 3 in the last few years of the time series were the most different between the models. However, data available to estimate recruitment in these years was limited. SSB estimates in the most recent years were similar in the two models, but the SS3 model resulted in larger estimates for SSB than those estimated by the 2011 model in most years (Figure 8). The fishery selectivity curves were nearly identical and thus cannot explain the differences in the trajectories of SSB (Figure 9). SS3 selectivity estimates resulted in lower proportions of older fish available to the survey compared to the 2011 model (Figure 10 & Figure 11). This may partially explain why SSB estimates in most years were higher for the SS3 results. Figure 12 shows observed and predicted survey biomass for the

2011 and SS3 models. The negative log likelihood for the survey biomass obtained with SS3 (-9.77) was substantially lower than that from the 2011 model (9.15), indicating that the SS3 model fit those data much better. This was apparent for the surveys conducted from 2006 to 2010 and from 1991 to 1995. In general, fits to age- and length-composition data are similar for both models (Figure 13-Figure 15), with some differences in predicted proportions-at-age for age 35-40+ fish (Figure 13) which resulted from differences in binning the age data. The 2011 model binned ages 35-39, while the data input to SS3 had separate age bins for each age up to age 40+; therefore, the age-composition data and expected values from the 2011 model look very large in Figure 13 for ages 35-40, as these are data points for two lumped age groups (35-39 and 40+), while the predicted age compositions for older ages from the SS3 model look small until age 40+ because an expected proportion (and a data point) exists for each older age that is younger than the plus group.

SUMMARY AND DISCUSSION OF DIFFERENCES BETWEEN THE SS3 MODEL AND 2011 MODEL

The differences between the configurations of the 2011 model and the SS3 model are:

(1) Both models used asymptotic selectivity curves, but the SS3 selectivity curves were parameterized with a double-normal with no descending limb (the standard deviation for the descending limb was set to a very high value), while the selectivity curves for the 2011 model were logistic. In addition, the 2011 model re-normalized the selectivity curves such that the largest selectivity occurs at 1. The asymptotic double-normal can approximate the logistic curve, but varied slightly. Numbers at age 3, SSB, and model fits for the SS3 model were similar to the 2011 model when fixing selectivity in the SS3 model to approximate the selectivity curves estimated in the 2011 model (Figure 16-Figure 21). However, SS3 selectivity estimates affected the fit to the data (Figure 9 & Figure 10) and the negative log likelihood for SS3's best model was –lnL = 1,282, while the negative log likelihood for the SS3 model with selectivity fixed to the curves most like those estimated in the 2011 model was –lnL = 2,670. SS3 does not have an option for normalizing the selectivity curves such that the greatest selectivity is always equal to 1, but the curve can be specified such that the peak value is at 1. SS3 runs conducted with a restriction that peak selectivity must equal 1 (and be asymptotic) estimated survey selectivity curves with selection occurring at smaller ages (e.g. Figure 22), leading to a poor fit to the survey data (Figure 23) and with a –lnL = 2586.

(2) The configuration of the likelihood components for early-period and main-period recruitment deviations differs between models. 2011 and SS3 model runs without recruitment deviations (recruitment deviations and weights for the recruitment likelihood components are set to 0) show that differences still exist between the models (Figure 24-Figure 28).

(3) SS3 population dynamics begin at age 0 and 2011 model dynamics begin at age 3. The SS3 model is given additional data, which consist of survey age-compositions for ages 0-2, separate age bins for ages 35-39 (rather than one lumped age bin), and length-compositions for lengths 0-17cm.

RESULTS: PROPOSED ALTERNATIVE SS3 MODELS

Table 3 shows the negative log likelihood components for each of the proposed alternative models (M1-M5) and the transitional SS3 model (M0). Figure 29- Figure 31 show a comparison of recruitment, recruitment deviations, and SSB for the proposed alternative models. All alternative models (M1-M5) had lower negative log likelihoods than the transitional SS3 model (M0). All models exhibited the same general trends over time in SSB and recruitment, but differences occurred in absolute numbers of recruits and SSB (Figure 29-Figure 31) with model M5 estimating the highest SSB and number of recruits and M3 the lowest SSB. Model M5, which estimates length-based, logistic, sex-specific selectivity, led to the lowest total negative log likelihood of any of the models, including the transitional SS3 model (Table 3). Model M5 did not fit the survey biomass index as well as the other models (Figure 43), but fit the age- and length composition data better (Table 3, Figure 44-Figure 48).

The selectivity curves for each model are shown in Figure 32-Figure 36. The length-based fishery selectivity curves that were estimated in models M1 and M4-M5 are similar to one another in each alternative model (Figure 32, Figure 35-Figure 36).

Models M4 and M5 led to the best total negative log likelihood values of the proposed models (-lnL = 1212.75 and 1183.51, respectively). Diagnostic plots for model M4 are shown in Figure 37-Figure 42 and the same plots are shown for model M5 in Figure 43-Figure 48. Model M4 was the best fit to the survey biomass index (-lnL = -12.38; Table 3; Figure 37), but did not fit the age- and length-composition data as well as model M5 (model M5 had a survey biomass index of -lnL = -6.74121; Table 3; Figure 43). The fits to the shallow-water survey length composition data were particularly poor for very young lengths for model M4; model M5 fits to the shallow-water survey length-composition data (where survey selectivity is length-based) were better.

An additional model run like model M4 was conducted, where the descending limb of the double-normal age-based selectivity curves were estimated; the resulting selectivity curves and other model results were identical to model M4, where age-based double-normal selectivity was forced to be asymptotic.

LITERATURE CITED

Abookire, A. A. and Macewicz, B. J. 2003. Latitudinal variation in reproductive biology and growth of female Dover sole (*Microstomus pacificus*) in the North Pacific, with emphasis on the Gulf of Alaska stock. Journal of Sea Research **50**: 187-197.

Methot, R. D. and C. R. Wetzel. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research **142**:86-99.

Stockhausen, W.T., M.E. Wilkins and M.H. Martin. 2011. 5. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 547-628. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510.

TABLES

Table 1. Symbols used in this document.

Symbol	Meaning
x	sex
a	age
f	fleet (fishery or survey)
t	time
$S_{f,x,a}$	Selectivity for fleet *f*, sex *x*, and age *a*
$N_{t,x,a}$	Numbers at age *a*, time *t*, and sex *s*
w_a	Weight at age *a*
$Z_{t,x,a}$	Total mortality at age *a*, sex *s*, and time *t*
timing	The timing of the survey during the year
$I_{t,f}$	Observed survey biomass at time *t* for fleet *f*
$SB_{t,f}$	Predicted survey biomass at time *t* for fleet *f*
$CV_{t,f}$	CV of observed survey biomass at time *t* for fleet *f*
$n_{t,x,f}$	Number of age-composition observations at time *t* for sex *x* and fleet *f*
$p_{t,x,f,a}$	Observed proportion at age *a*, time *t*, fleet *f*, and sex *x*
$\hat{p}_{t,x,f,a}$	Predicted proportion at age *a*, time *t*, fleet *f*, and sex *x*
$n_{2,t,x,f}$	Number of length-composition observations at time *t* for sex *x* and fleet *f*
$p_{t,x,f,l}$	Observed proportion at length *l*, time *t*, fleet *f*, and sex *x*
$\hat{p}_{t,x,f,l}$	Predicted proportion at length *l*, time *t*, fleet *f*, and sex *x*
$\tilde{R}_t$	Estimated mean recruitment in year *t*
σ_R	Recruitment CV (specified in SS3 only)
b_t	Bias adjustment factor at time *t* (specified in SS3 only)
C_t^{obs}	Observed catch at time *t*
$\hat{C}_t$	Predicted catch at time *t*
$\sigma_{t,f}$	Standard error of catch at time *t* for fleet *f* (specified for SS3 only)

Table 2. Likelihood components used in the 2011 and SS3 models. Numbers in the component column are likelihood component weightings for: (SS3, 2011 Model).

Component	SS3	2011 Model
Survey biomass $(SB_{t,f})$ equation	$\sum_x \sum_a S_{f,x,a} N_{t,x,a} w_a \exp(-\text{timing}(Z_{t,x,a}))$	$\sum_x \sum_a S_{f,x,a} N_{t,x,a} w_a$
Survey biomass likelihood (1,1)	$\sum_{t \in survey\ f} \frac{(\ln(\mathrm{I}_{t,f}) - \ln(SB_{t,f}))^2}{2\ln(CV_{t,f}^2 + 1)}$	As for SS3
Age composition (1, 1)	$\sum_t \sum_x \sum_a n_{t,x,f} p_{t,x,f,a} \ln\left(\frac{p_{t,x,f,a}}{\hat{p}_{t,x,f,a}}\right)$	As for SS3
Length Composition (0.5, 0.5)	$\sum_t \sum_x \sum_l n_{2,t,x,f} p_{t,x,f,1} \ln\left(\frac{p_{t,x,f,l}}{\hat{p}_{t,x,f,l}}\right)$	As for SS3
Main period recruits (1,1)	$\frac{1}{2}\left(\sum_{t=1984}^{2008}\left(\frac{\tilde{R}_t^2}{\sigma_R^2} + b_t \ln(\sigma_R^2)\right)\right)$ (sum to 0 constraint)	$\sum_{t=1947}^{2008} \tilde{R}_t^2$ (sum to 0 constraint)
Early period recruits (1,2)	$\frac{1}{2}\left(\sum_{t=1947}^{1983}\left(\frac{\tilde{R}_t^2}{\sigma_R^2} + \ln(\sigma_R^2)\right)\right)$	$\sum_{t=1947}^{1983} \tilde{R}_t^2$ (sum to 0 constraint)
Late period recruits (1,3)	$\frac{1}{2}\left(\sum_{t=2009}^{2011}\left(\frac{\tilde{R}_t^2}{\sigma_R^2} + \ln(\sigma_R^2)\right)\right)$	$\sum_{t=2009}^{2011} \tilde{R}_t^2$ (sum to 0 constraint)
Catch (30,30)	$\sum_t \frac{(\ln(\mathrm{C}_t^{obs}) - \ln(\hat{\mathrm{C}}_t))^2}{2\sigma_{t,f}^2}$	$\sum_t (\ln(\mathrm{C}_t^{obs}) - \ln(\hat{\mathrm{C}}_t))^2$

Table 3. Components of the negative log(likelihood) for each alternative proposed SS3 model. M0-M5 are the alternative model descriptors, which are described in full in the section “Analytic Approach: Proposed Alternative SS3 Models” on page 7. The “Total” likelihoods marked “but add’l component” include an additional likelihood component for initial equilibrium catch and therefore the likelihoods cannot be compared directly to those alternative models where a component for initial equilibrium catch was not estimated. However, the contribution of the initial equilibrium catch likelihood component to the total negative log(likelihood) is very small in each case.

Likelihood component	M0	M1	M2	M3	M4	M5
Total (not always comparable to the transitional model)	1281.56	1216.90	1275.38 (but add'l component)	1281.94	1212.75 (but add'l component)	1183.51 (but add'l comonent)
Initial Equilibrium Catch	NA	NA	0.0020	NA	0.0014	8.96E-05
Survey Biomass	-9.7695	-10.8824	-11.2050	-9.5223	-12.3772	-6.74121
Length Composition	847.0220	775.6090	846.8010	847.07	778.3310	773.034
Age Composition	445.6960	455.3270	442.6560	445.81	451.7830	429.217
Recruitment	-1.5099	-3.1827	-2.9911	-1.5326	-5.0088	-12.0118

FIGURES

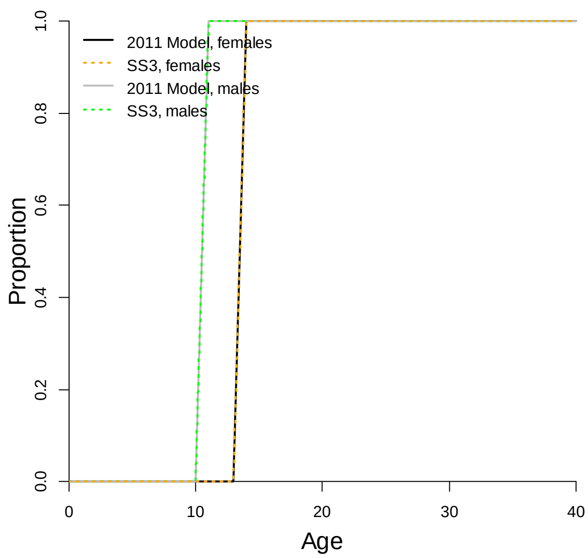


Figure 1. Fishery selectivity for Dover sole used in deterministic runs to match population dynamics between the 2011 and SS3 models. Selectivity curves are fixed at MLEs for fishery selectivity from the 2011 model. The SS3 selectivity curves pictured were created using a double-normal selectivity functional form with no descending limb; the 2011 model selectivity curves are logistic.

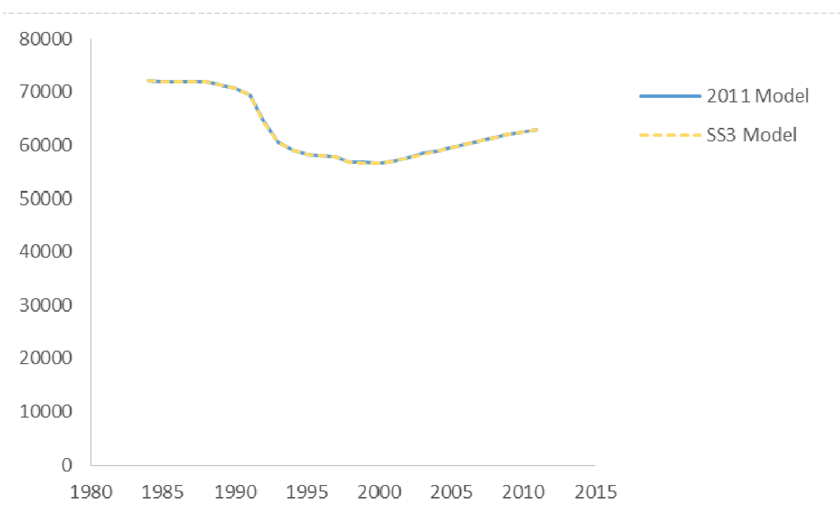


Figure 2. Spawning stock biomass for a deterministic run of the 2011 and SS3 models with parameters fixed at the MLEs for the 2011 Dover sole model with Dover sole catch history and no recruitment deviations. Fishery selectivity curves for the models were forced to match as closely as possible (Figure 1).

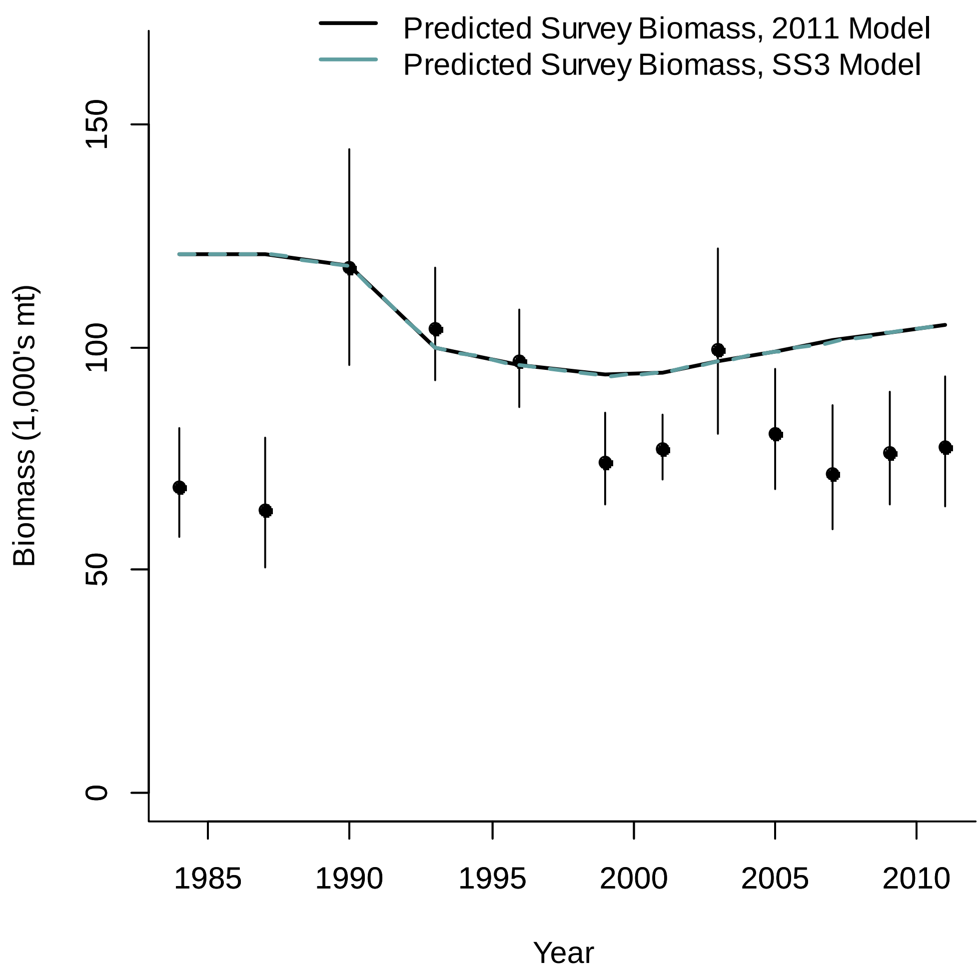


Figure 3. Survey biomass for the 2011 model (black solid line) and the SS3 model (blue dashed line) for a deterministic run with no estimation, parameters fixed at the same values in both models, and fishery and survey selectivity curves in both models fixed to the curve shown in Figure 1.

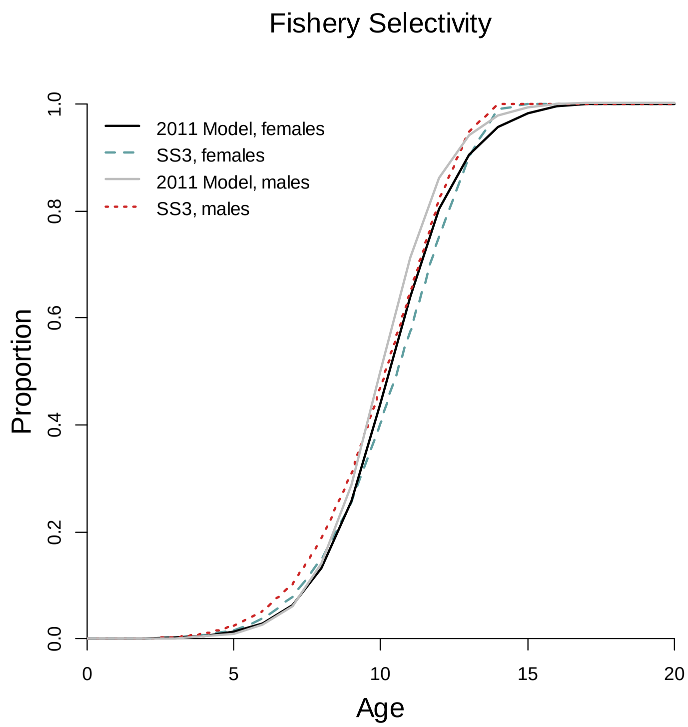


Figure 4. Example SS3 double-normal selectivity curves that fail to match the 2011 model's logistic fishery selectivity curves exactly (the standard deviation of the descending limb of the selectivity curves was fixed at a large value to create an asymptotic curve).

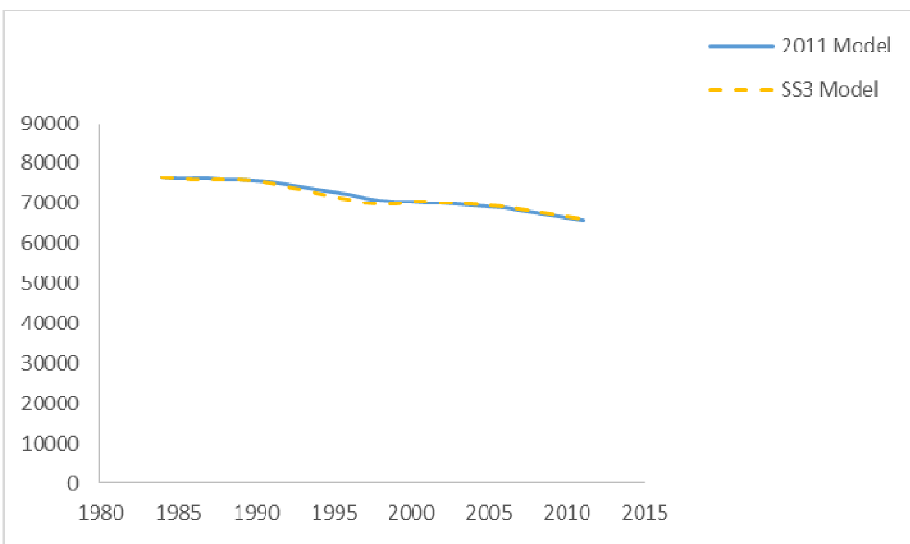


Figure 5. Spawning stock biomass for a deterministic run of the 2011 and SS3 models with parameters in both models fixed at the same values, using flathead sole catch history with no recruitment deviations. Fishery selectivity curves for the models were forced to match as closely as possible, but are not an exact match (Figure 4).

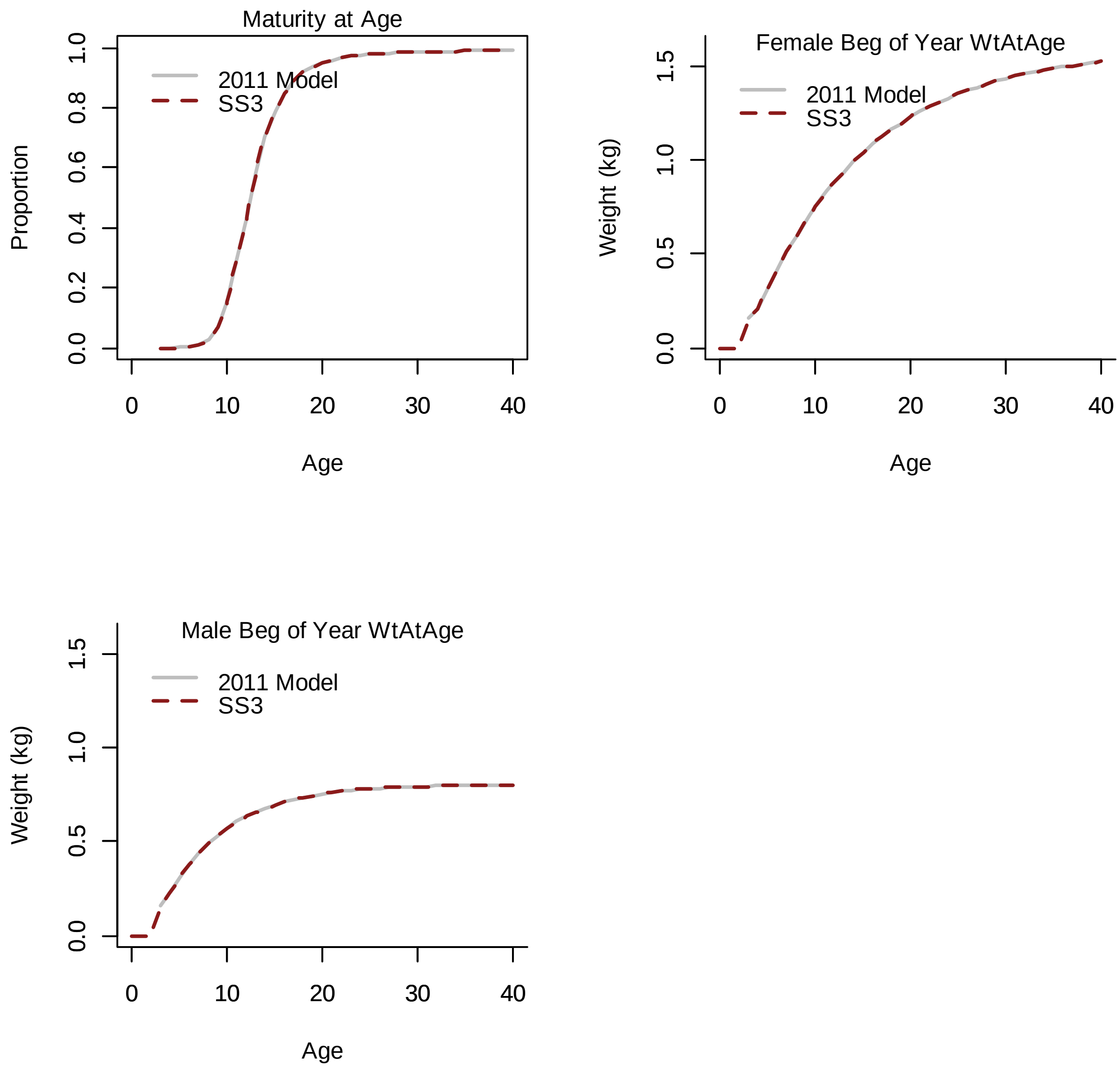


Figure 6. Maturity and weight-at-age for males and females (also used as mid-year weight at age) for the 2011 model and an equivalent SS3 model. The lines match perfectly because both models use empirical vectors for each of the three relationships.

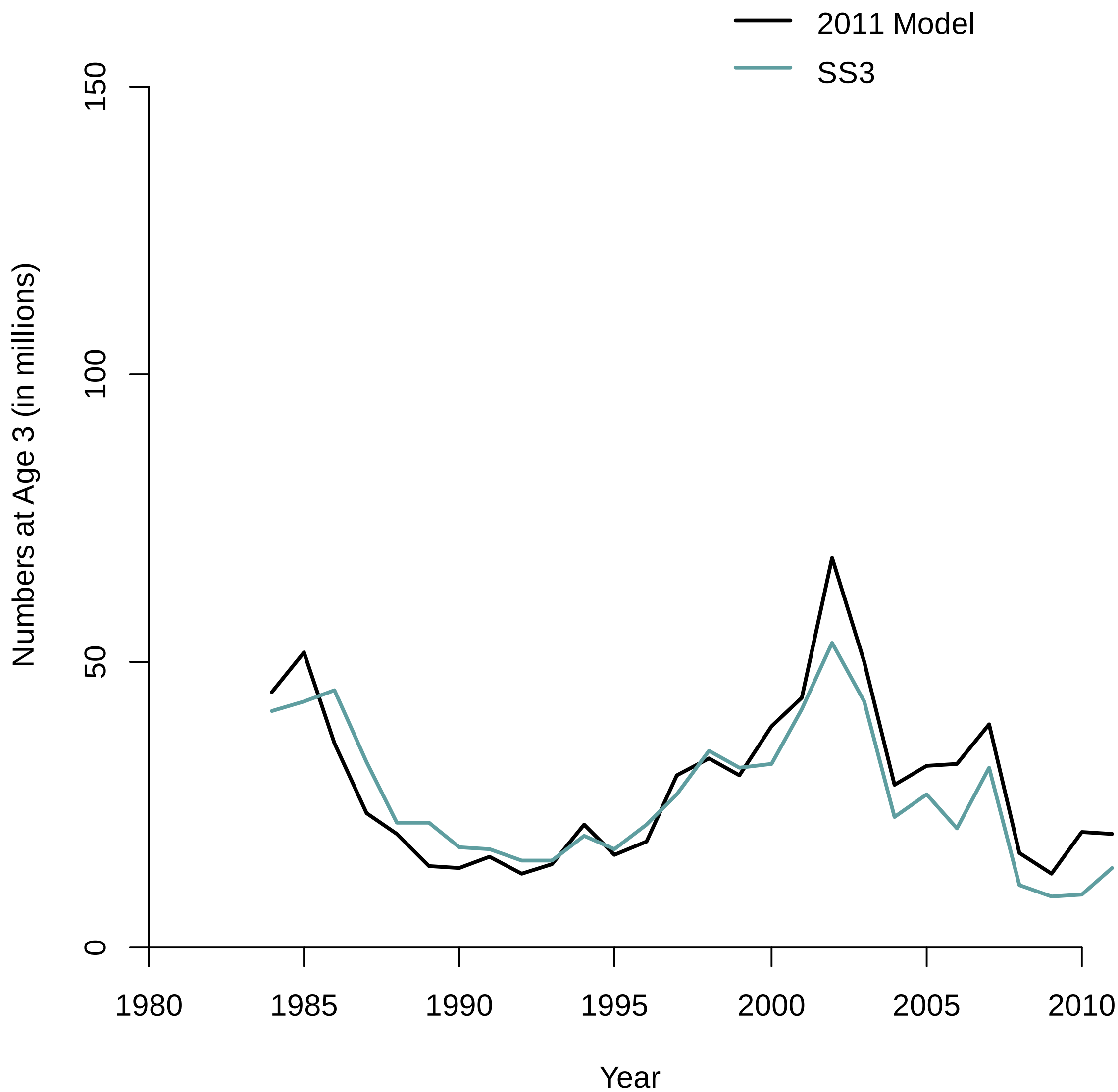


Figure 7. Numbers at age 3 for the 2011 model (black line) and an equivalent SS3 run (blue line). Both models estimate the log of mean recruitment, recruitment deviations for 1984-2011, an early period of recruitment deviations starting in 1964, fishing mortality rates, and asymptotic selectivity parameters (logistic for the 2011 model and double-normal for SS3). Survey data for ages 0-2 and lengths 0-18cm are included in the SS3 model, but not the 2011 model.

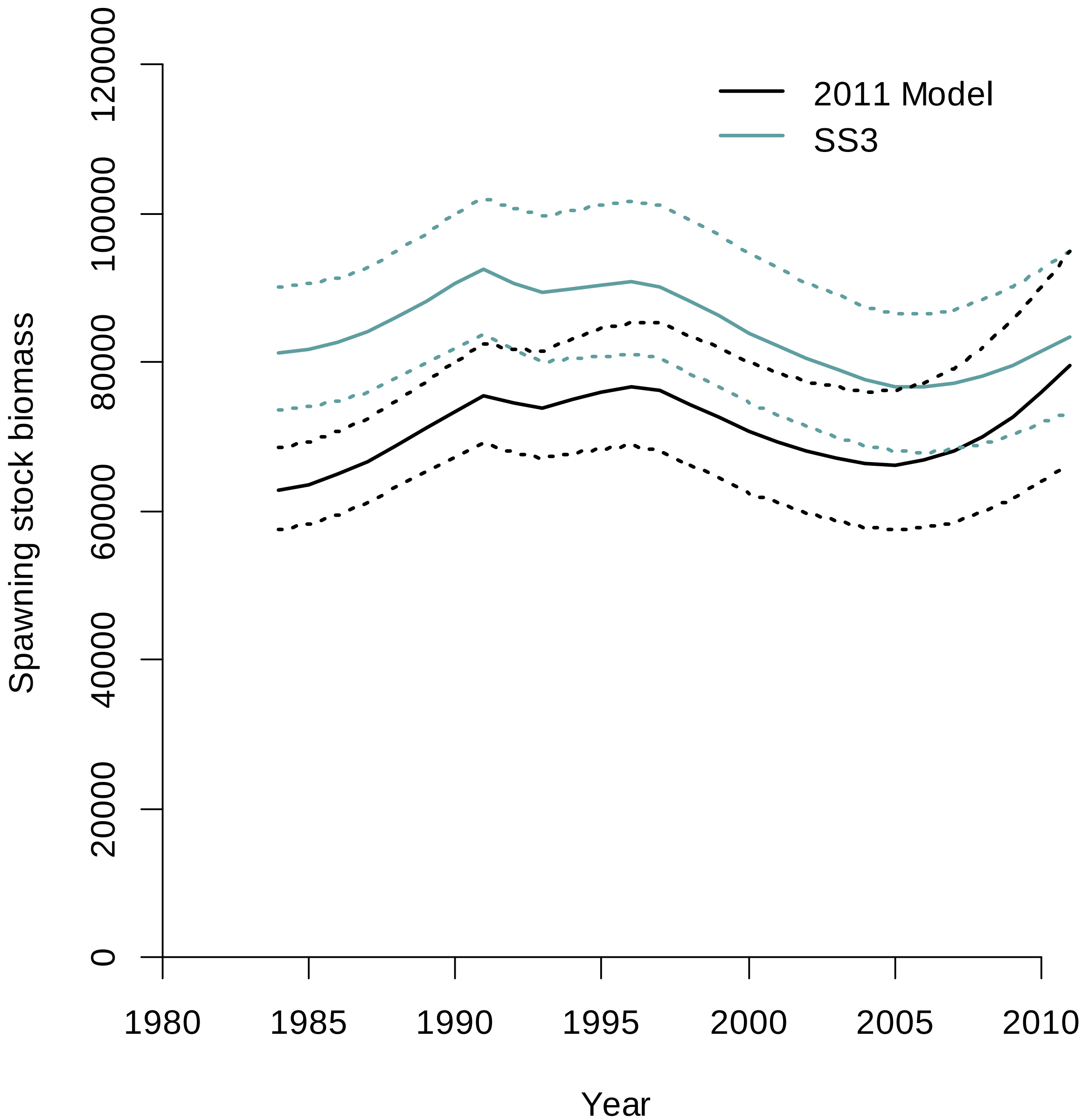


Figure 8. Spawning stock biomass (solid lines) and asymptotic 95% confidence intervals (dotted lines) for the 2011 model (black lines) and SS3 (blue lines) for an equivalent SS3 model. Both models estimate the log of mean recruitment, recruitment deviations for 1984-2011, an early period of recruitment deviations starting in 1964, fishing mortality rates, and asymptotic selectivity parameters (logistic for the 2011 model and double-normal for SS3). Survey data for ages 0-2 and lengths 0-18cm are included in the SS3 model, but not the 2011 model.

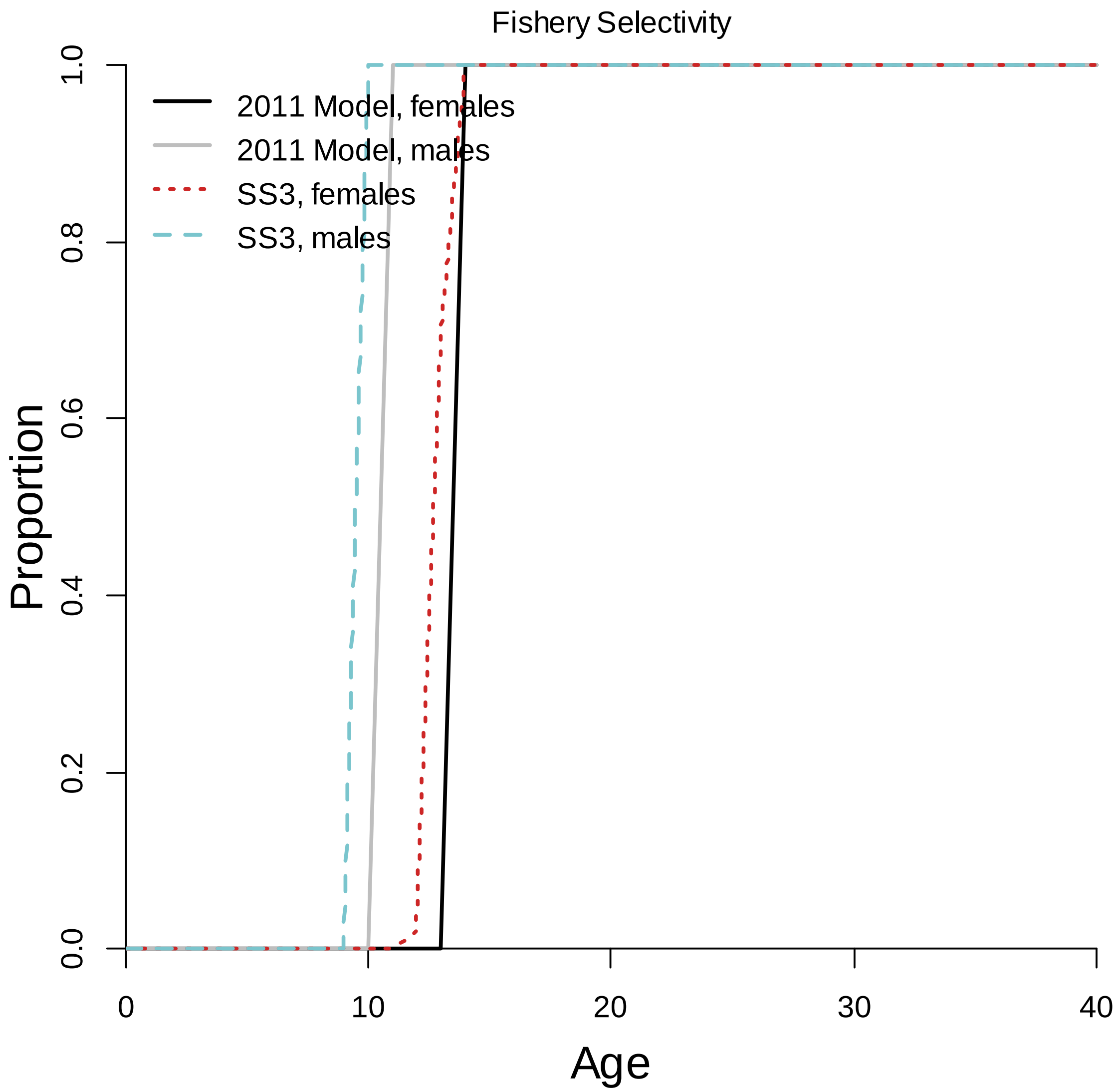


Figure 9. Fishery selectivity for the 2011 model (solid lines) and an equivalent SS3 model run (dotted and dashed lines).

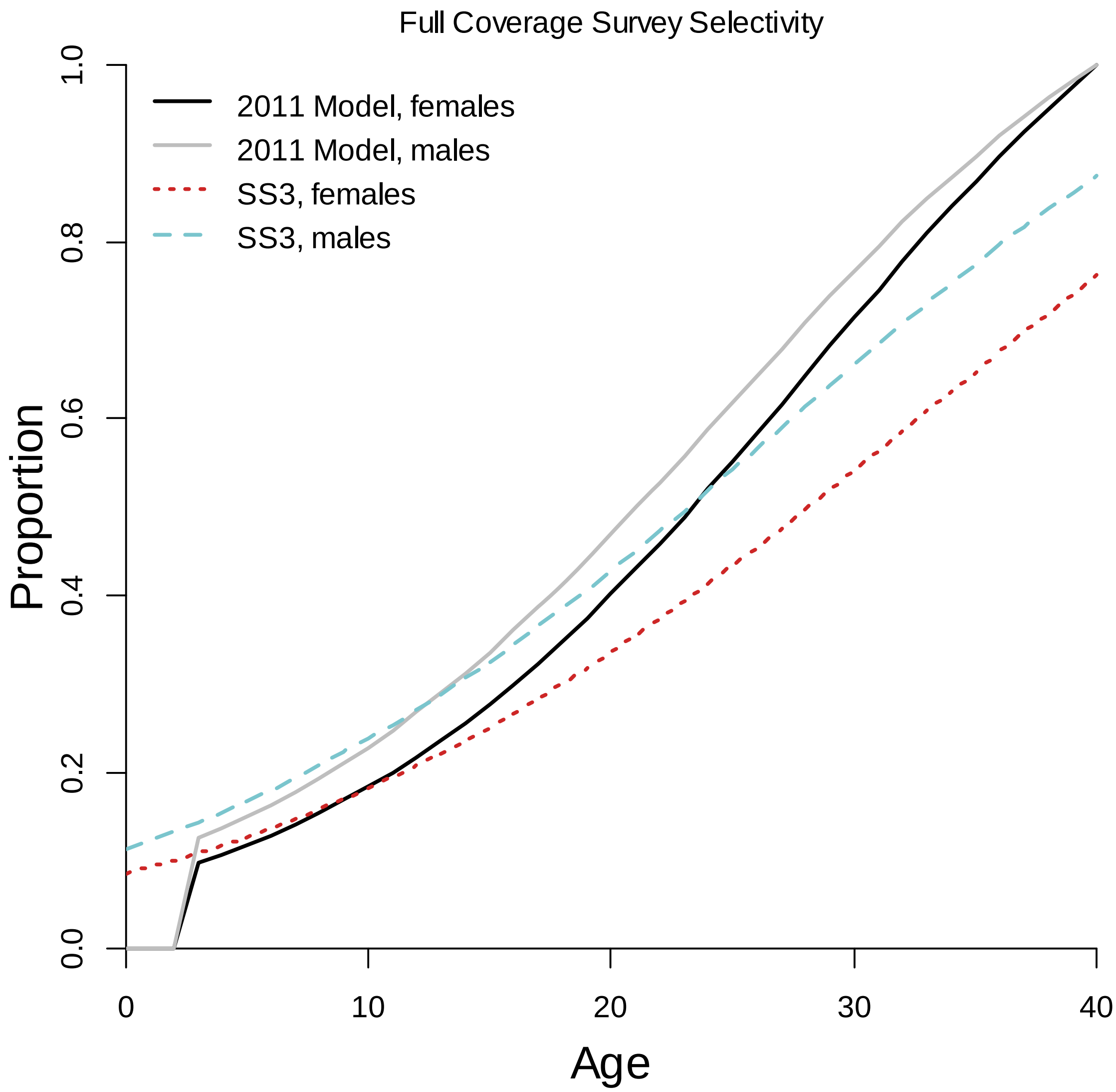


Figure 10. Survey selectivity for the 2011 model (solid lines) and an equivalent SS3 model run (dotted and dashed lines) for years with fuller survey coverage.

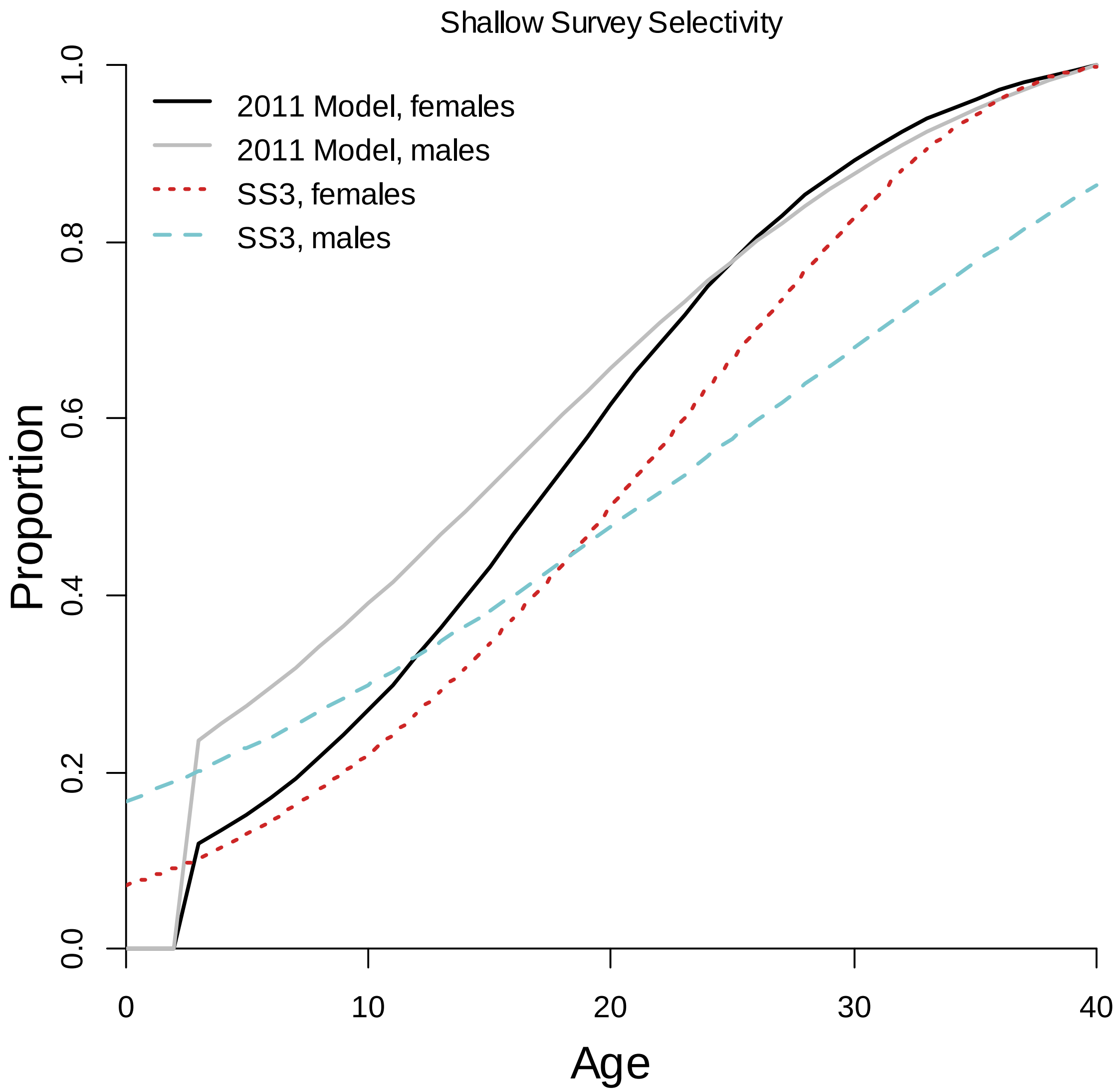


Figure 11. Survey selectivity for the 2011 model (solid lines) and an equivalent SS3 model run (dotted and dashed lines) for years with only shallower water survey coverage.

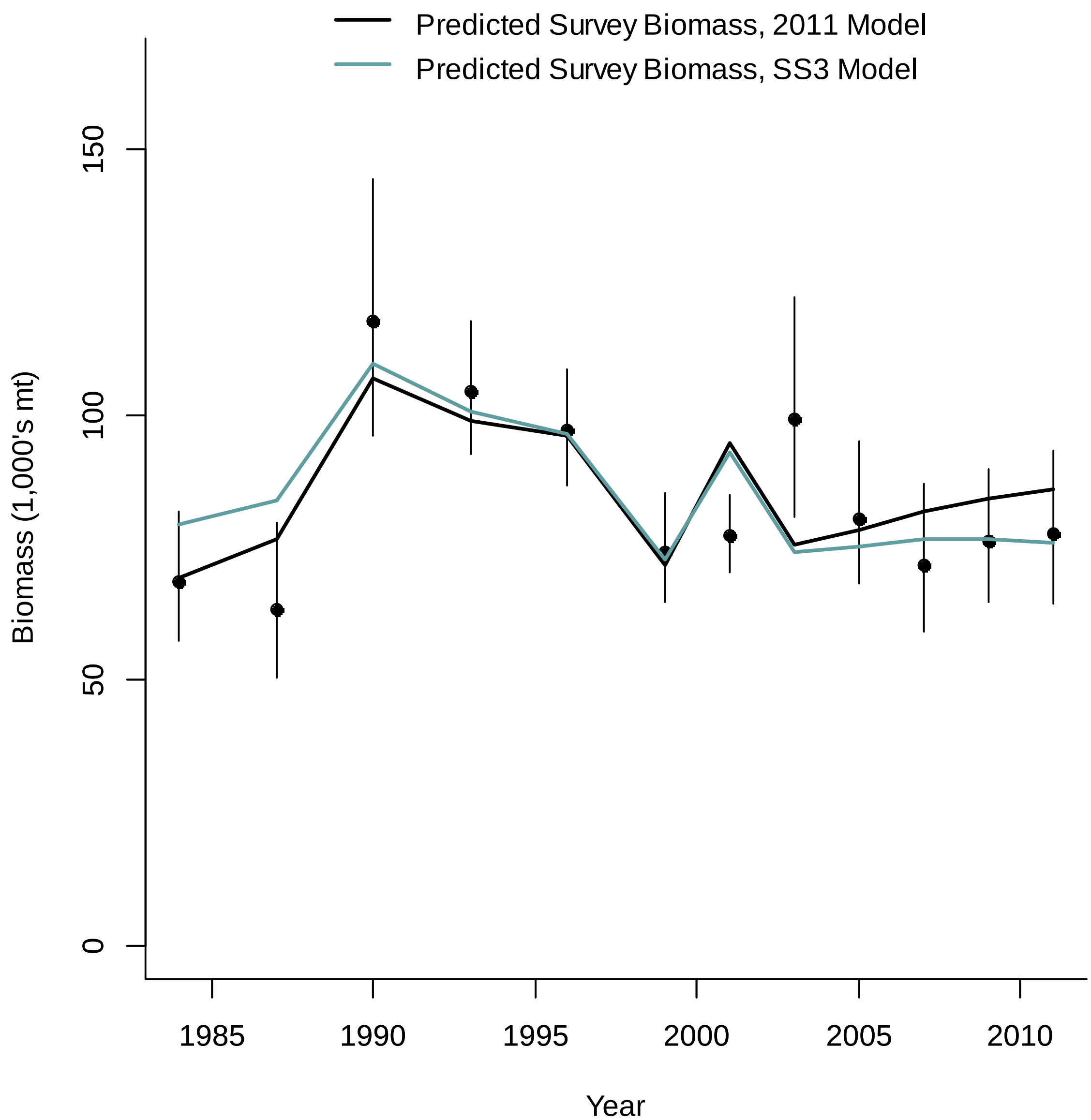
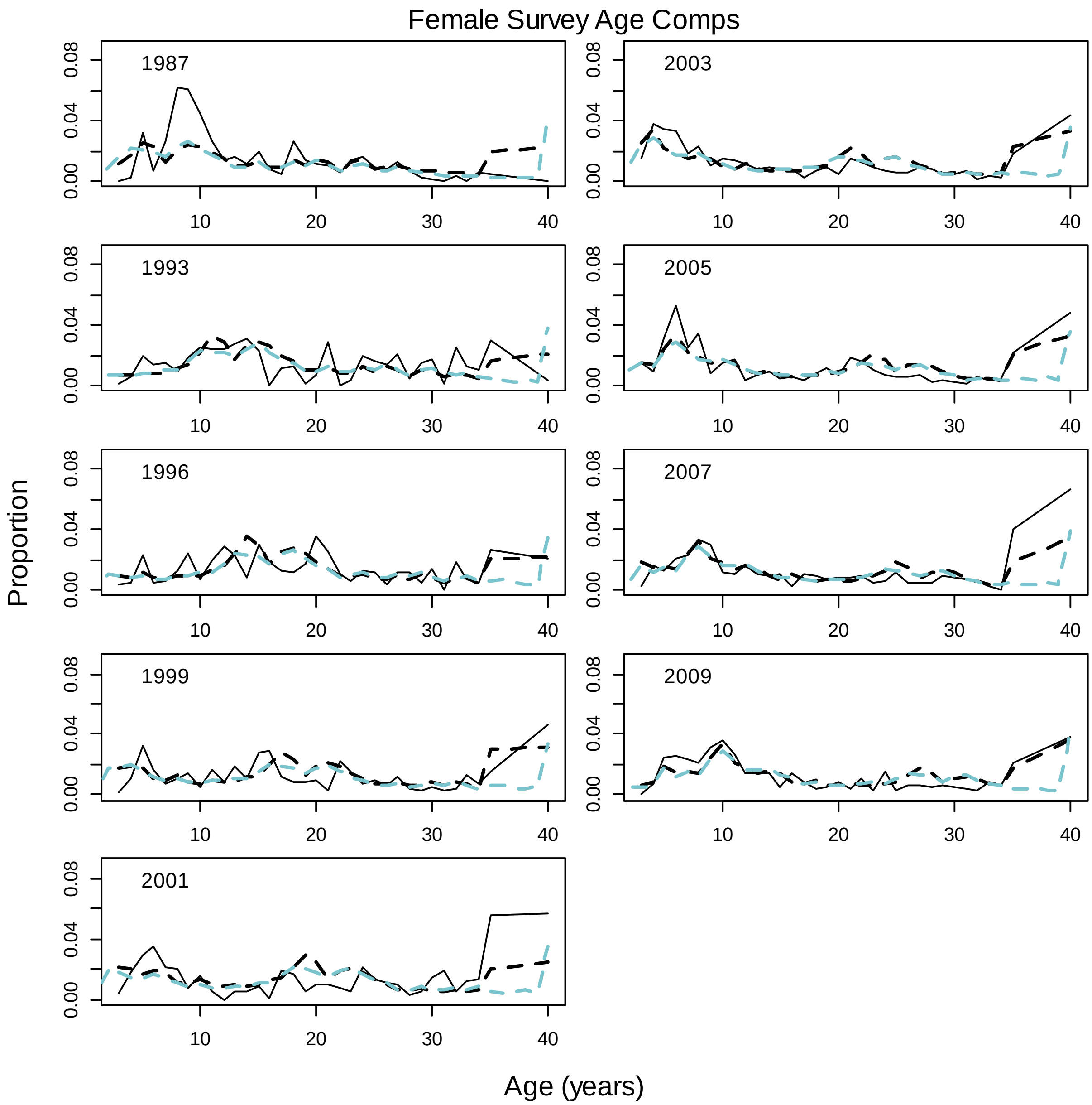


Figure 12. Observed survey biomass (black dots) with 95% asymptotic confidence intervals (vertical black lines) and predicted survey biomass from the 2011 model (black line) and an equivalent SS3 model (blue line).


Female Survey Age Comps
1987
2003
1993
2005
1996
2007
1999
2009
2001
Proportion
Age (years)
0.00
0.04
0.08
10
20
30
40

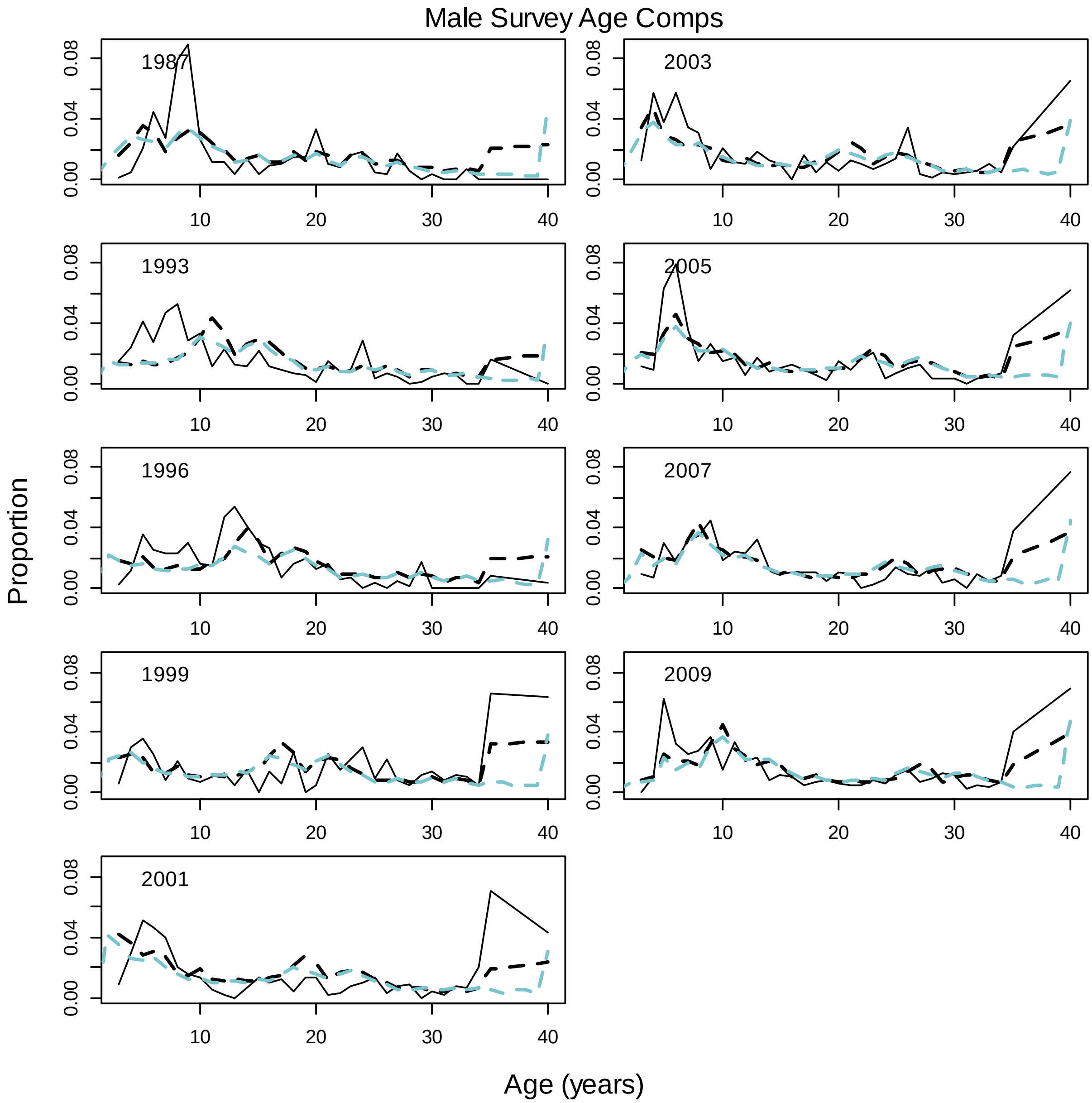
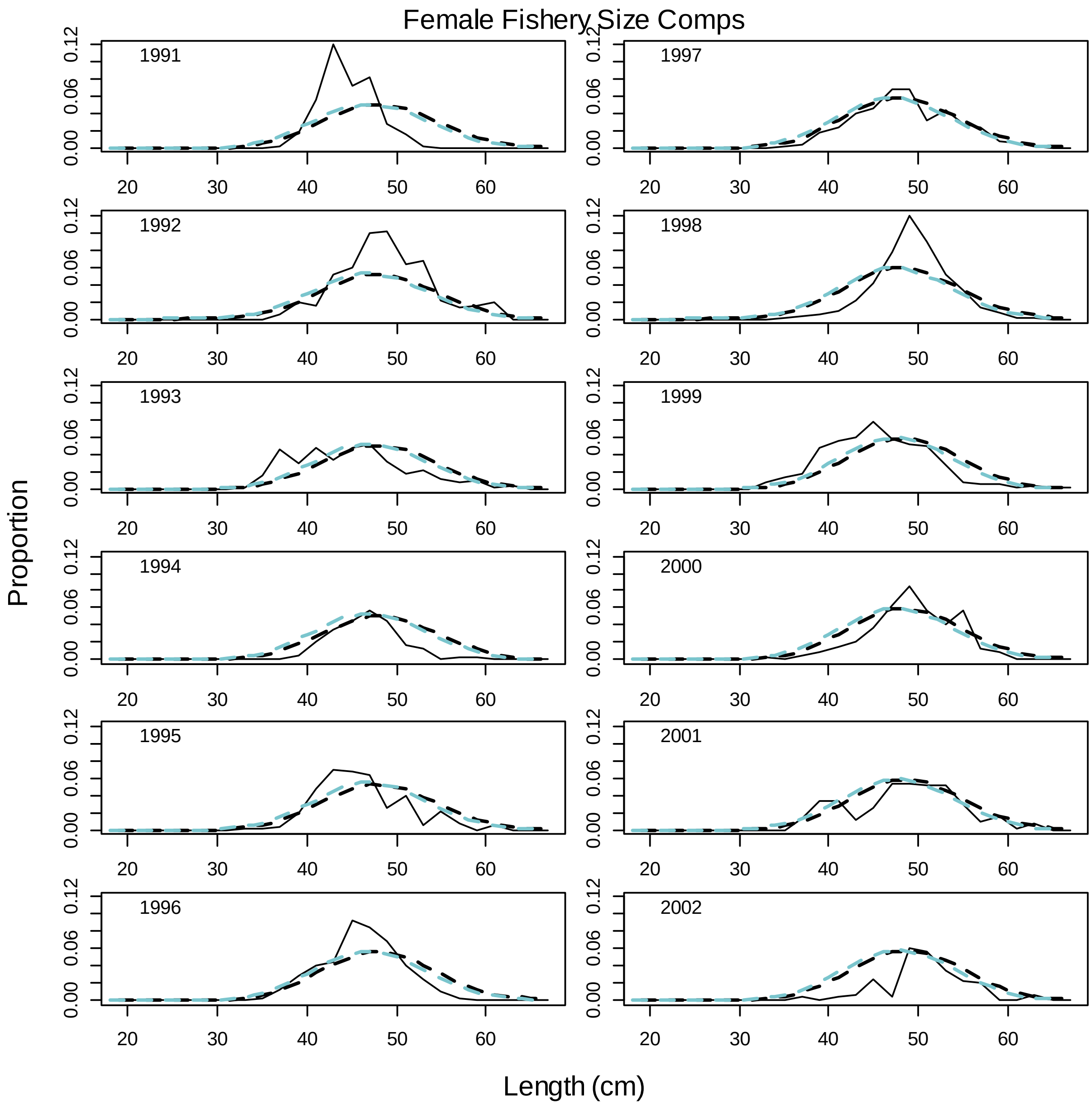
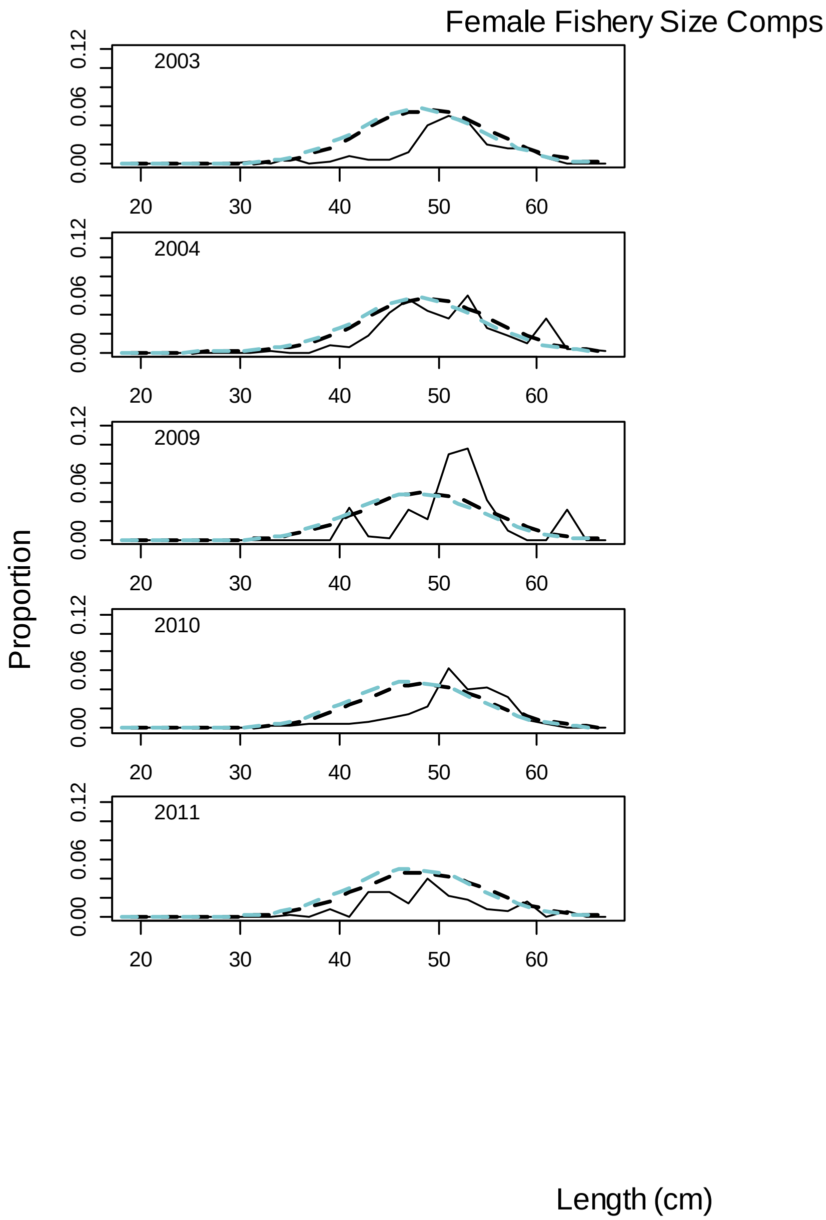
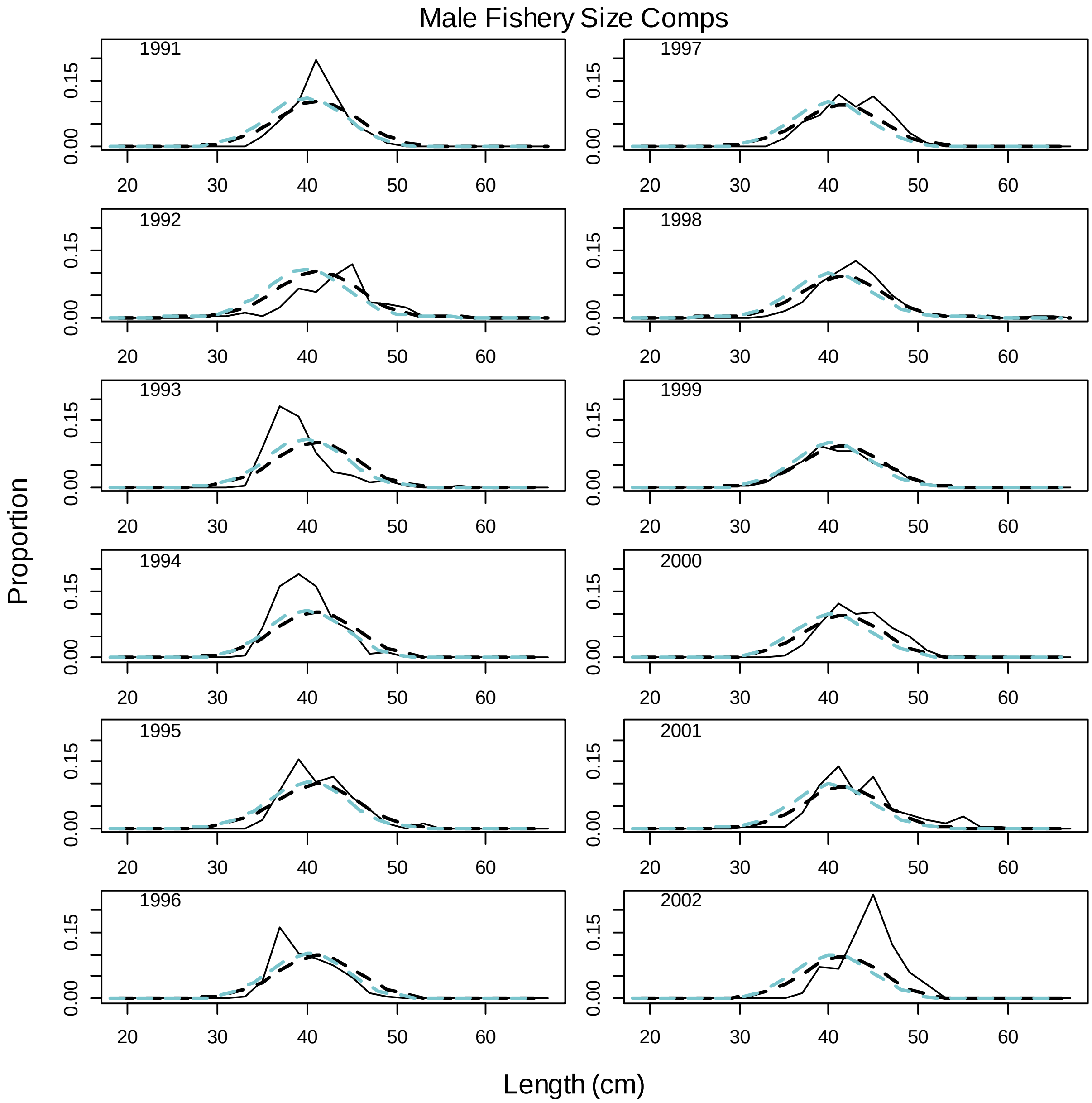


Figure 13. Observed (2011 model; solid black lines) and predicted (dashed lines) survey proportions-at-age for the 2011 model (dashed black lines) and an equivalent SS3 model run (dashed blue lines) for females (first panel) and males (second panel). The SS3 model included data for age 0-2 individuals and ages 35-40 were each separate age bins, while the 2011 model included data from ages 3-40 with an age 35 bin that included ages 35-39. Expectations for the SS3 model therefore do not match those from the 2011 model (or the 2011 data) for ages 35-39 and 0-2.


Female Fishery Size Comps
Proportion
Length (cm)
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
0.00
0.06
0.12
20
30
40
50
60


Female Fishery Size Comps
Proportion
Length (cm)
2003
2004
2009
2010
2011
0.00 0.06 0.12
20 30 40 50 60


Male Fishery Size Comps
Proportion
Length (cm)
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
0.00
0.15
20
30
40
50
60

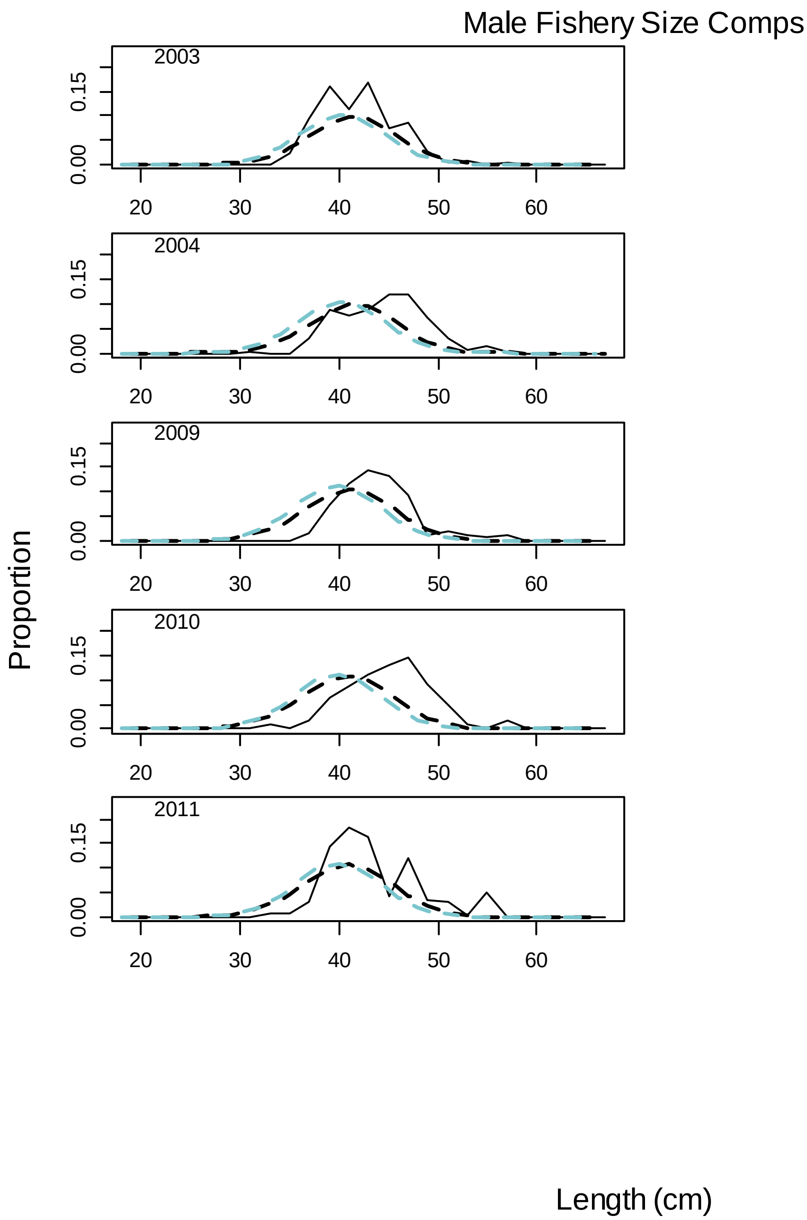
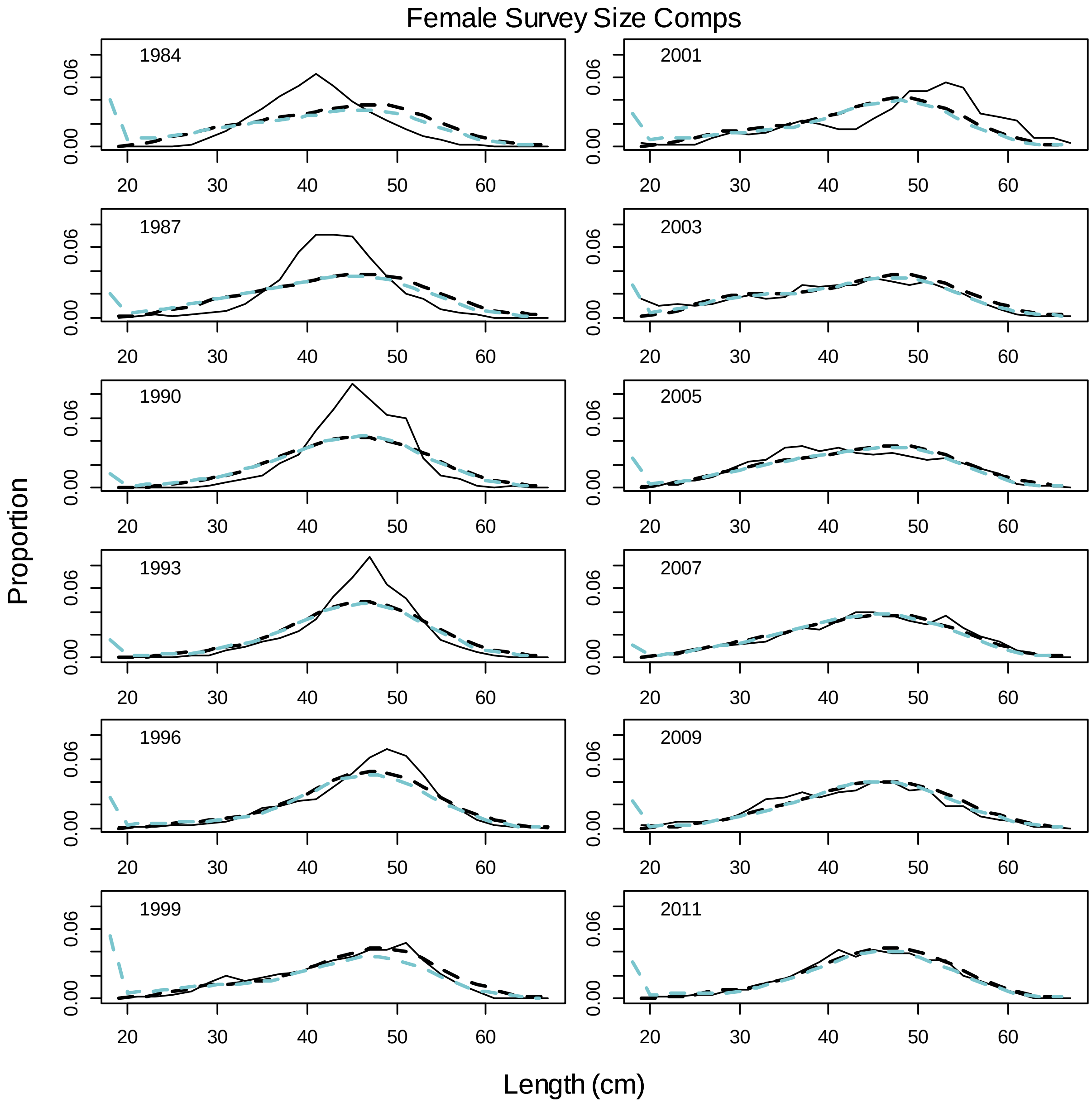


Figure 14. Observed (solid black lines) and predicted (dashed lines) fishery proportions-at-length for the 2011 model (dashed black lines) and an equivalent SS3 model run (dashed blue lines) for females (first set of panels) and males (second set of panels).


Female Survey Size Comps
Proportion
Length (cm)
1984
2001
1987
2003
1990
2005
1993
2007
1996
2009
1999
2011
0.00
0.06
20
30
40
50
60

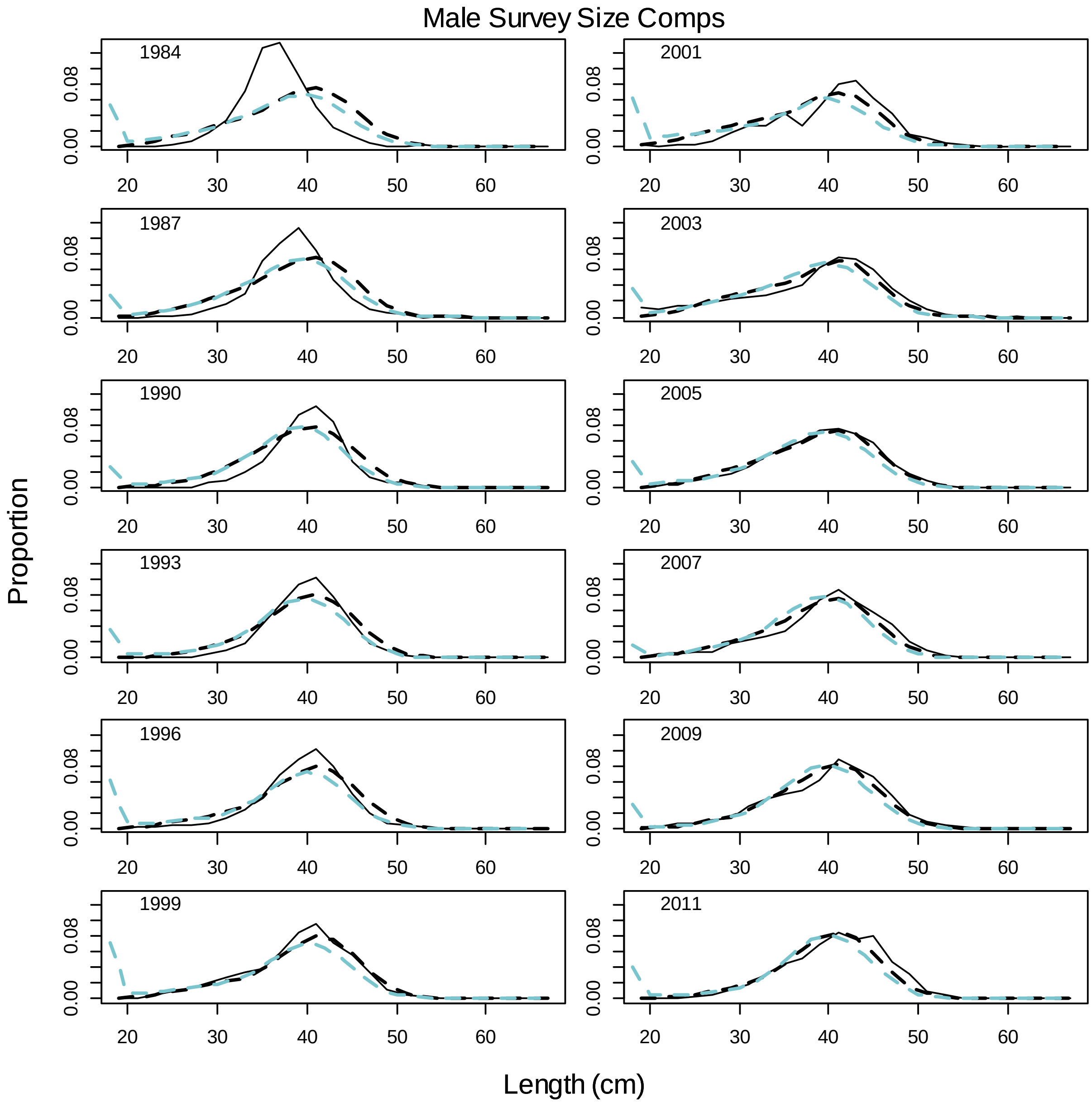


Figure 15. Observed (solid black lines) and predicted (dashed lines) survey proportions-at-length for the 2011 model (dashed black lines) and an equivalent SS3 model run (dashed blue lines) for females (first set of panels) and males (second set of panels).

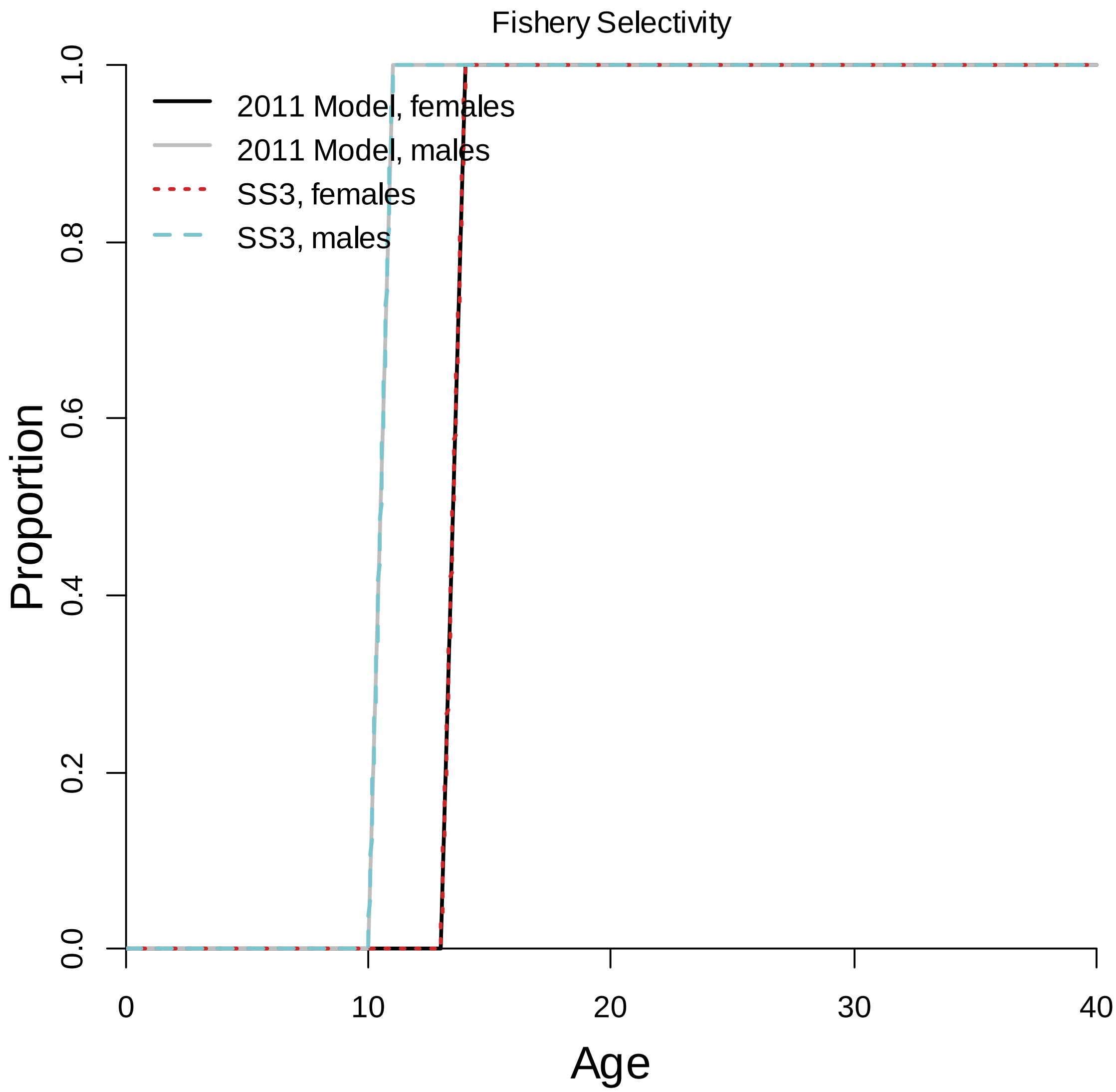


Figure 16. Fishery selectivity for the 2011 models and an SS3 model with selectivity fixed to be as similar as possible to the selectivity curves estimated in the 2011 model.

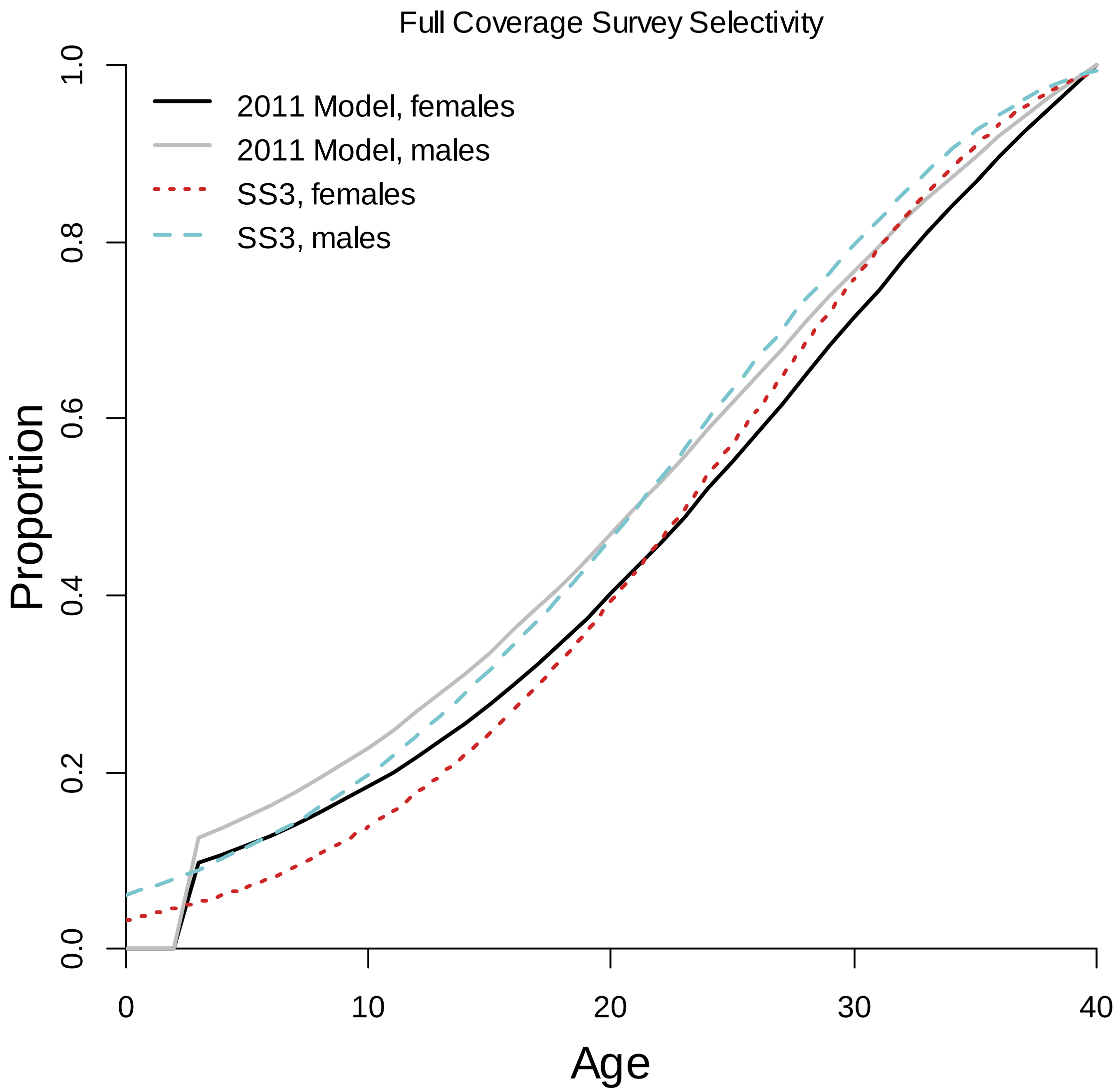


Figure 17. Full coverage survey selectivity for the 2011 models and an SS3 model with selectivity fixed to be as similar as possible to the selectivity curves estimated in the 2011 model.

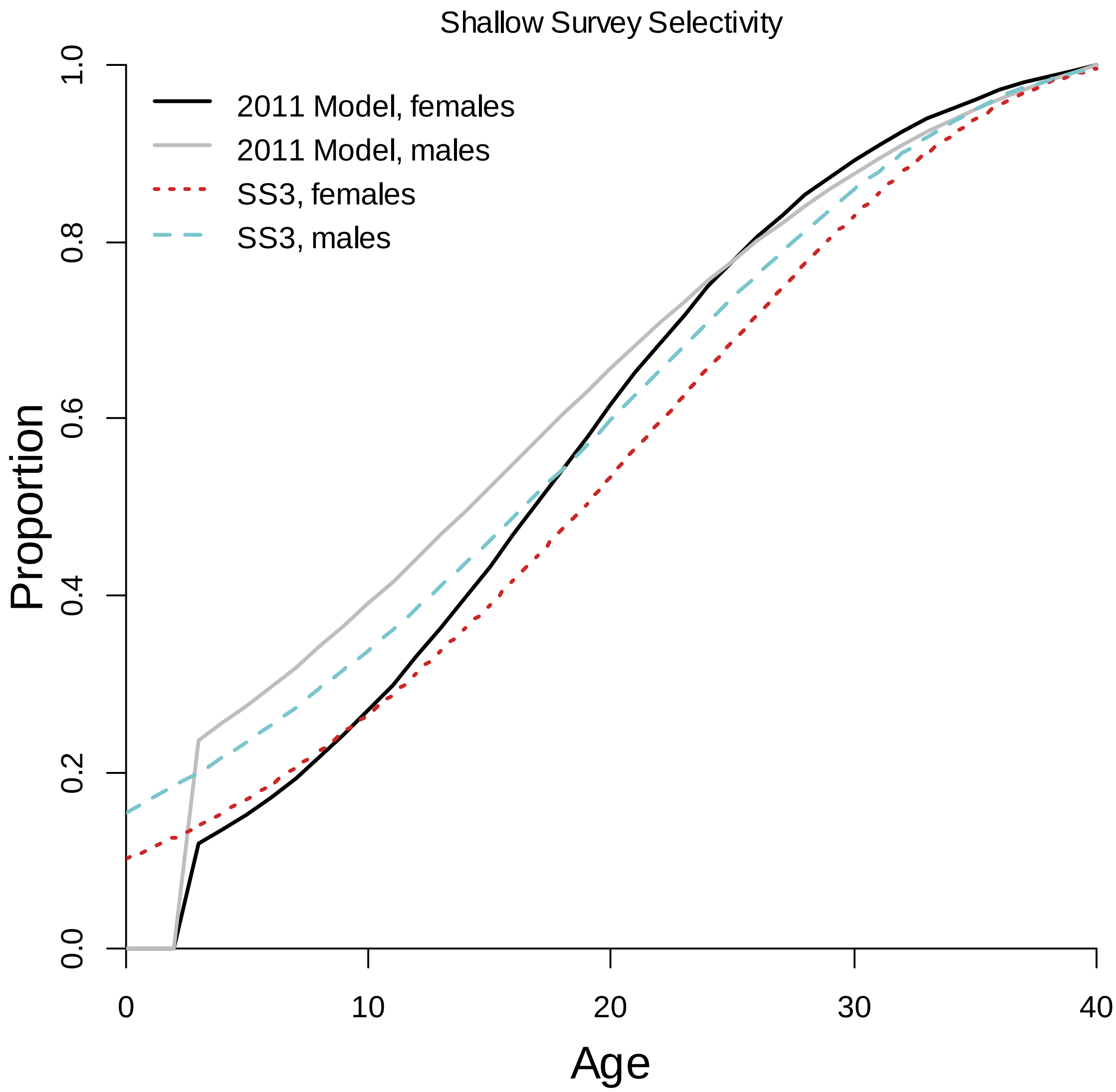


Figure 18. Shallow water survey selectivity for the 2011 models and an SS3 model with selectivity fixed to be as similar as possible to the selectivity curves estimated in the 2011 model.

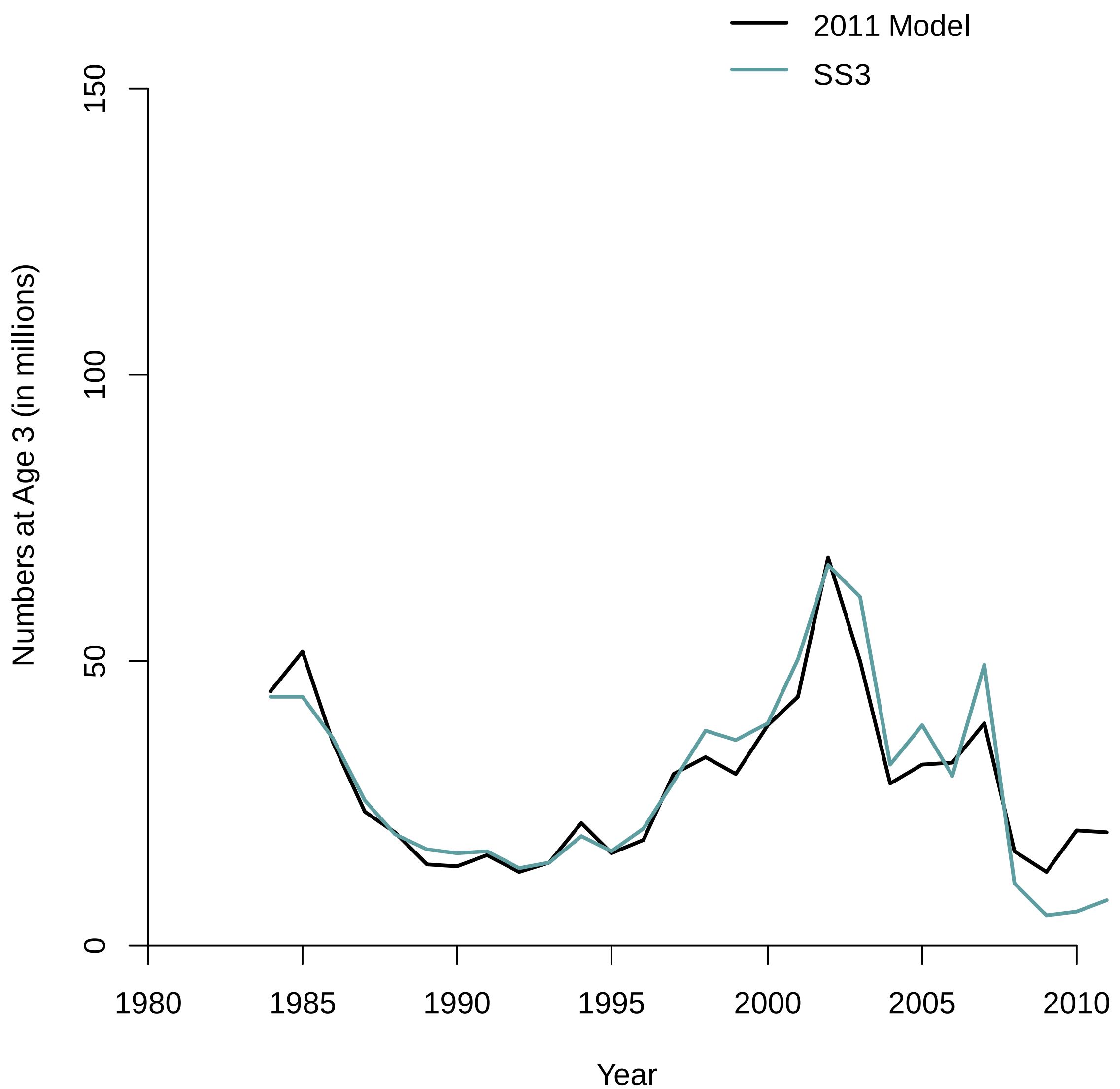


Figure 19. Numbers at age 3 for the 2011 models and an SS3 model with selectivity fixed to be as similar as possible to the selectivity curves estimated in the 2011 model.

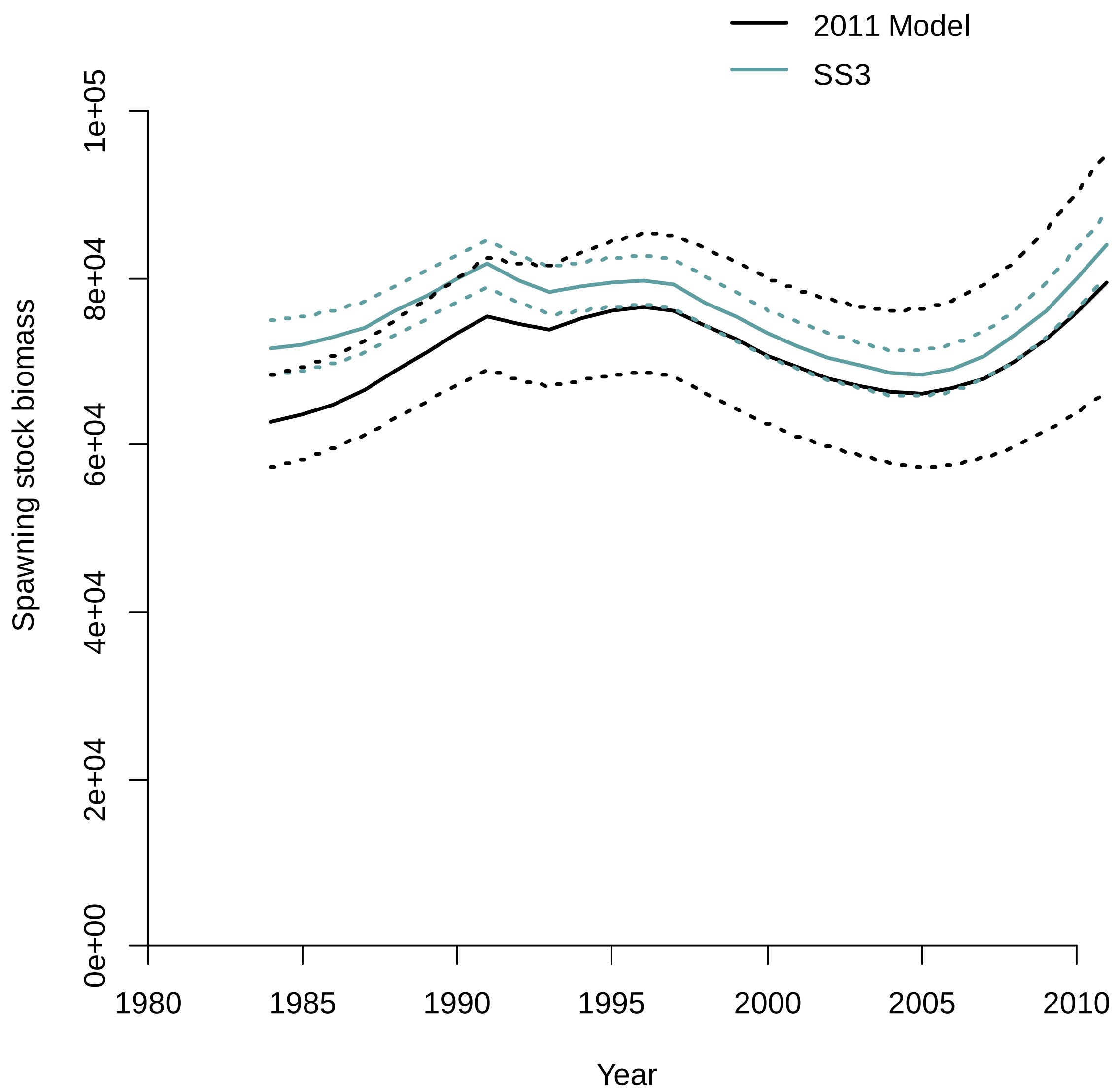


Figure 20. Spawning stock biomass (solid lines) and 95% asymptotic confidence intervals (dotted lines) for the 2011 models (black lines) and an SS3 model (blue lines) with selectivity fixed to be as similar as possible to the selectivity curves estimated in the 2011 model.

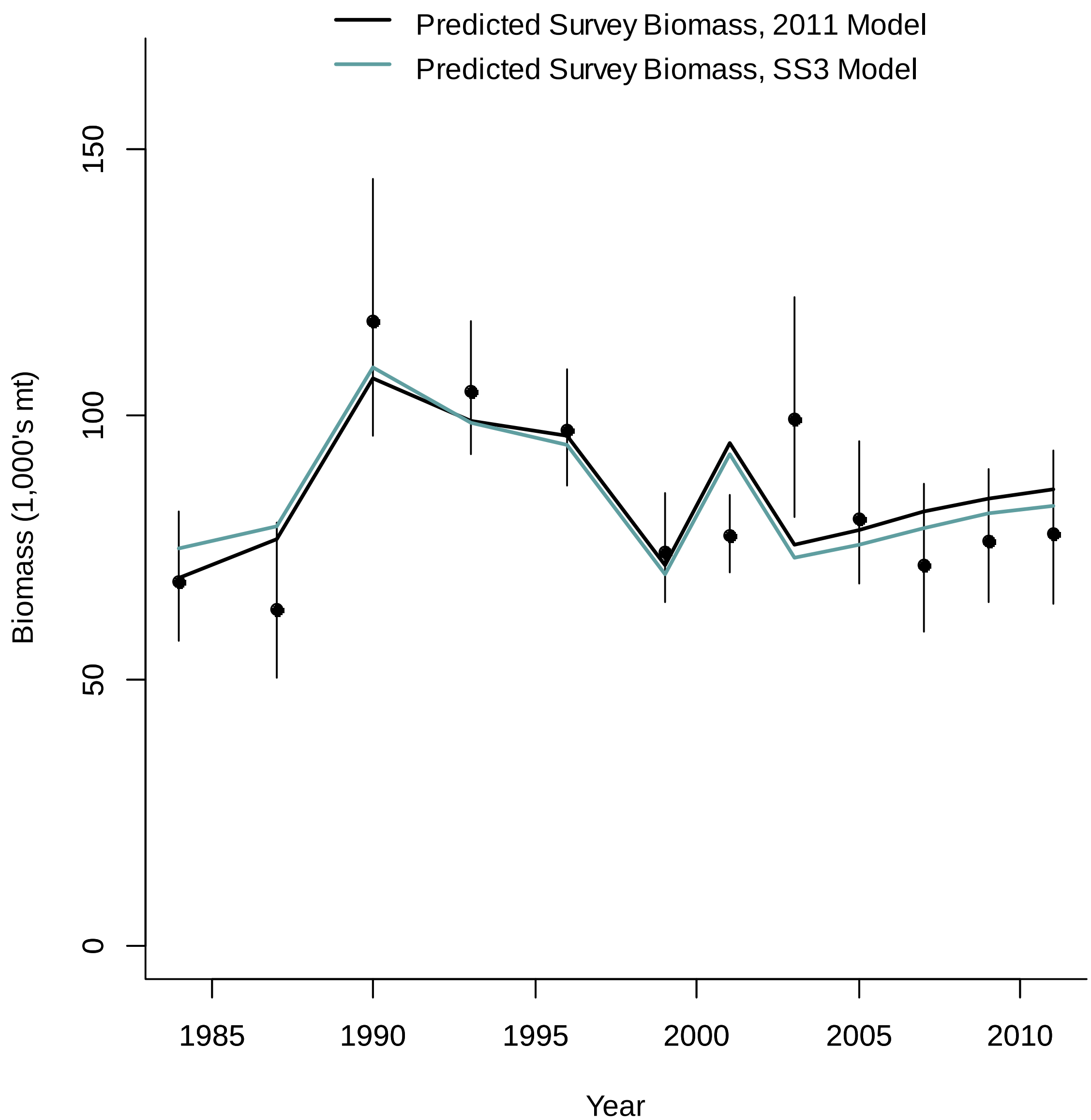


Figure 21. Observed survey biomass (black dots) with 95% asymptotic confidence intervals (vertical black lines) and predicted survey biomass for the 2011 models (black lines) and an SS3 model (blue lines) with selectivity fixed to be as similar as possible to the selectivity curves estimated in the 2011 model.

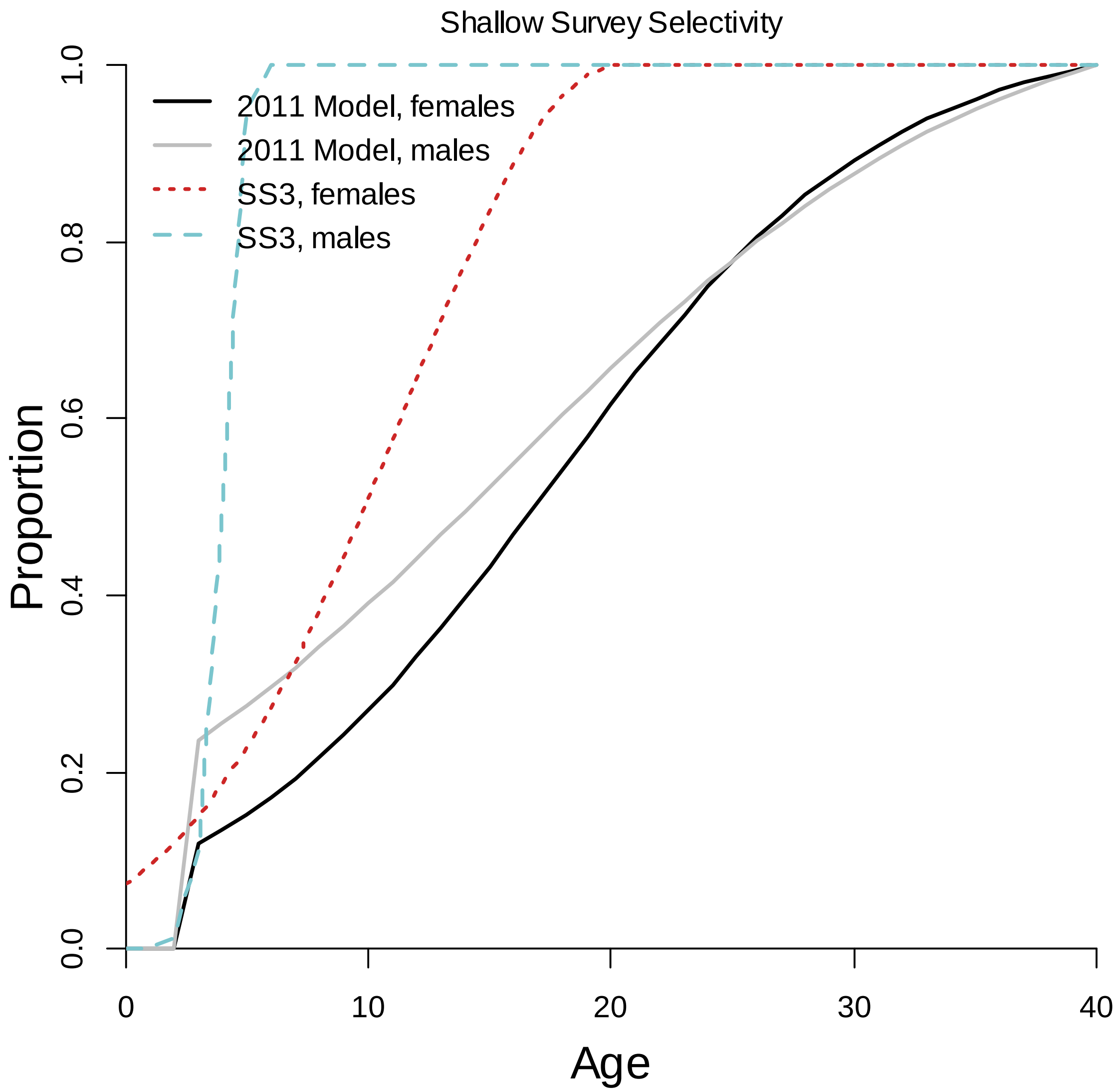


Figure 22. Shallow water survey selectivity for the 2011 model and an SS3 model with estimation of selectivity restricted such that it must reach 1 at or below age 40.

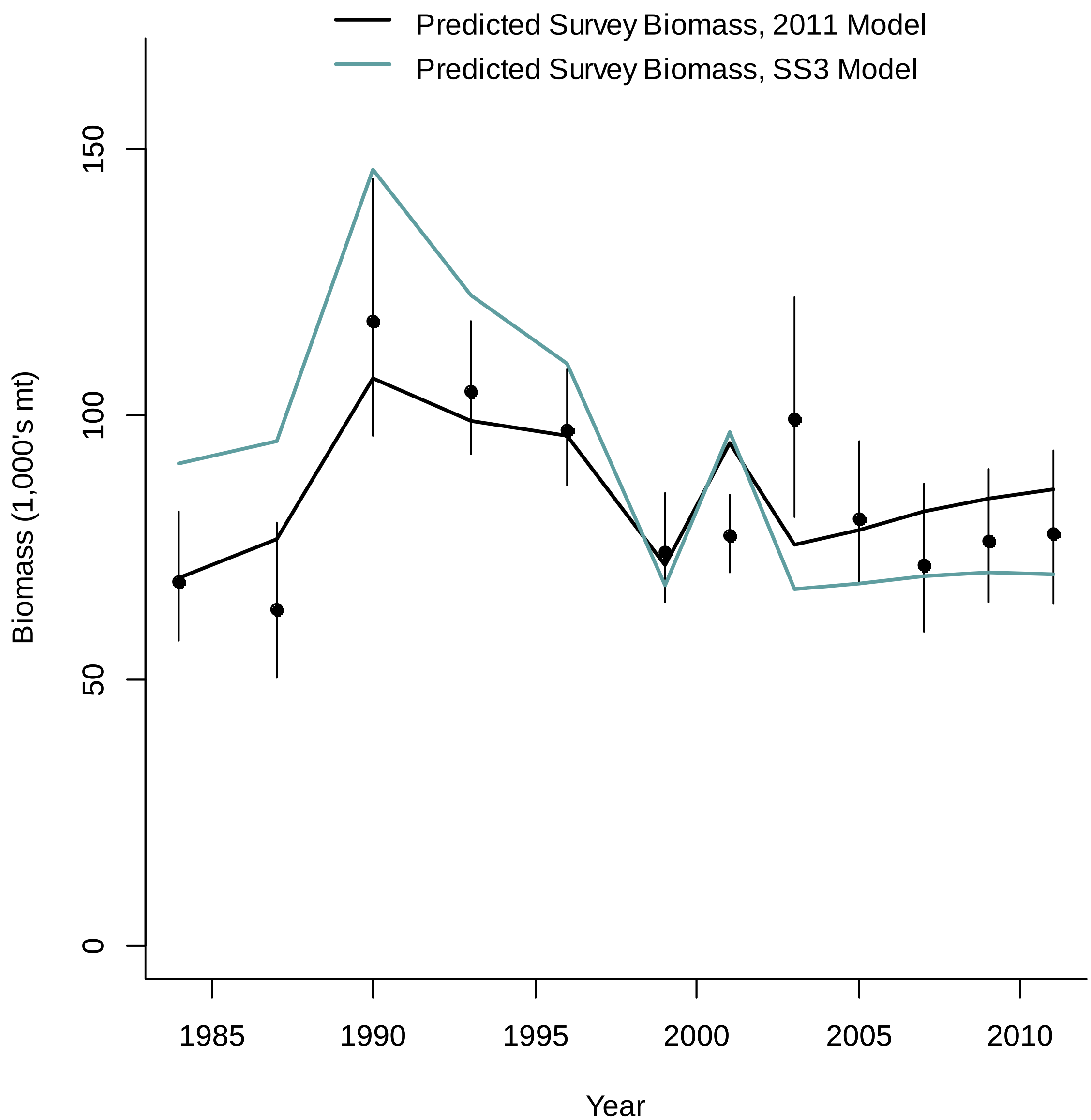


Figure 23. Observed survey biomass (black dots) with 95% asymptotic confidence intervals (vertical black lines) and predicted survey biomass for the 2011 models (black lines) and an SS3 model (blue lines) with estimation of selectivity restricted such that it must reach 1 at or below age 40.

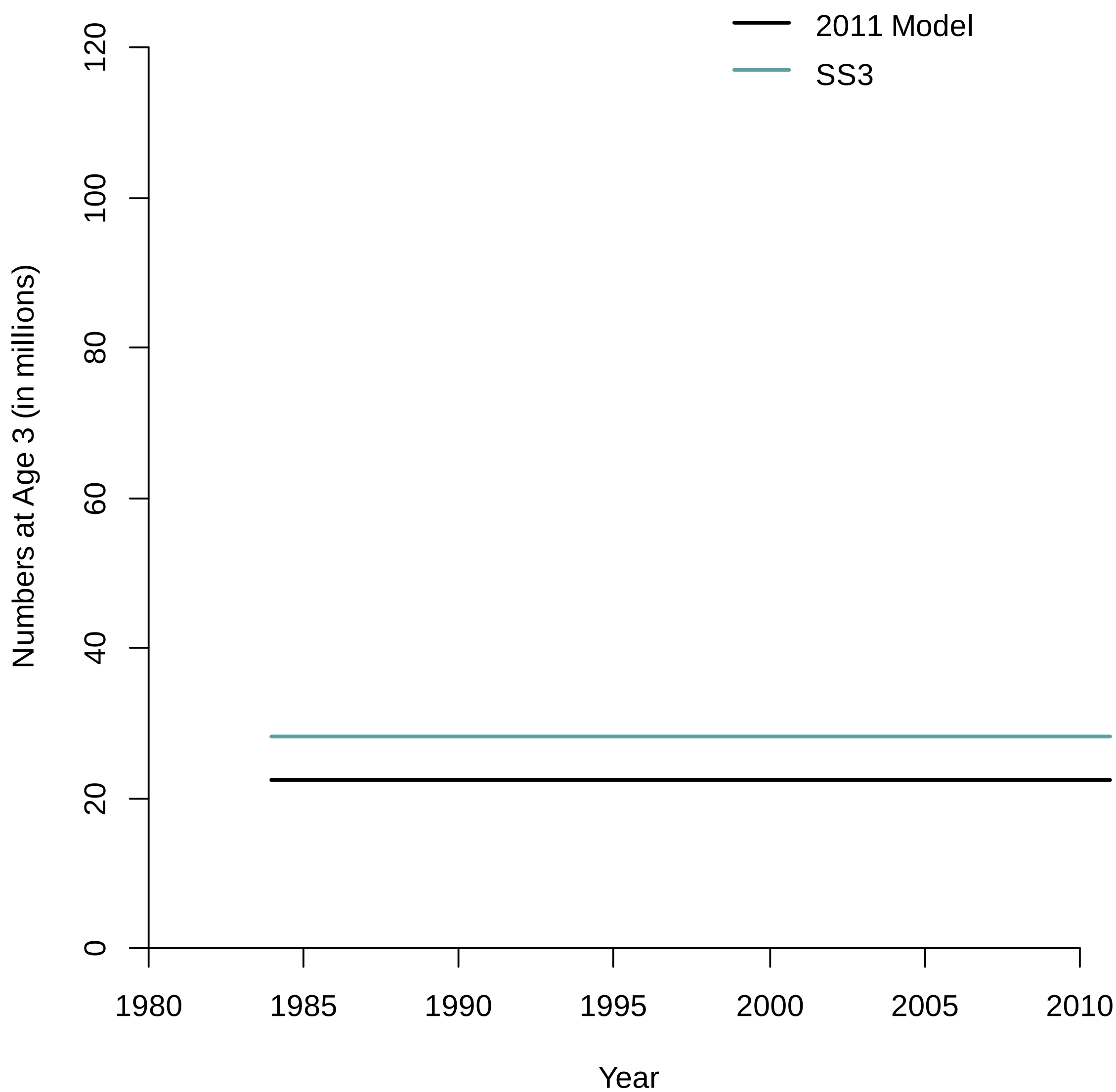


Figure 24. Numbers at age 3 for runs of the SS3 and 2011 models without estimation of recruitment deviations.

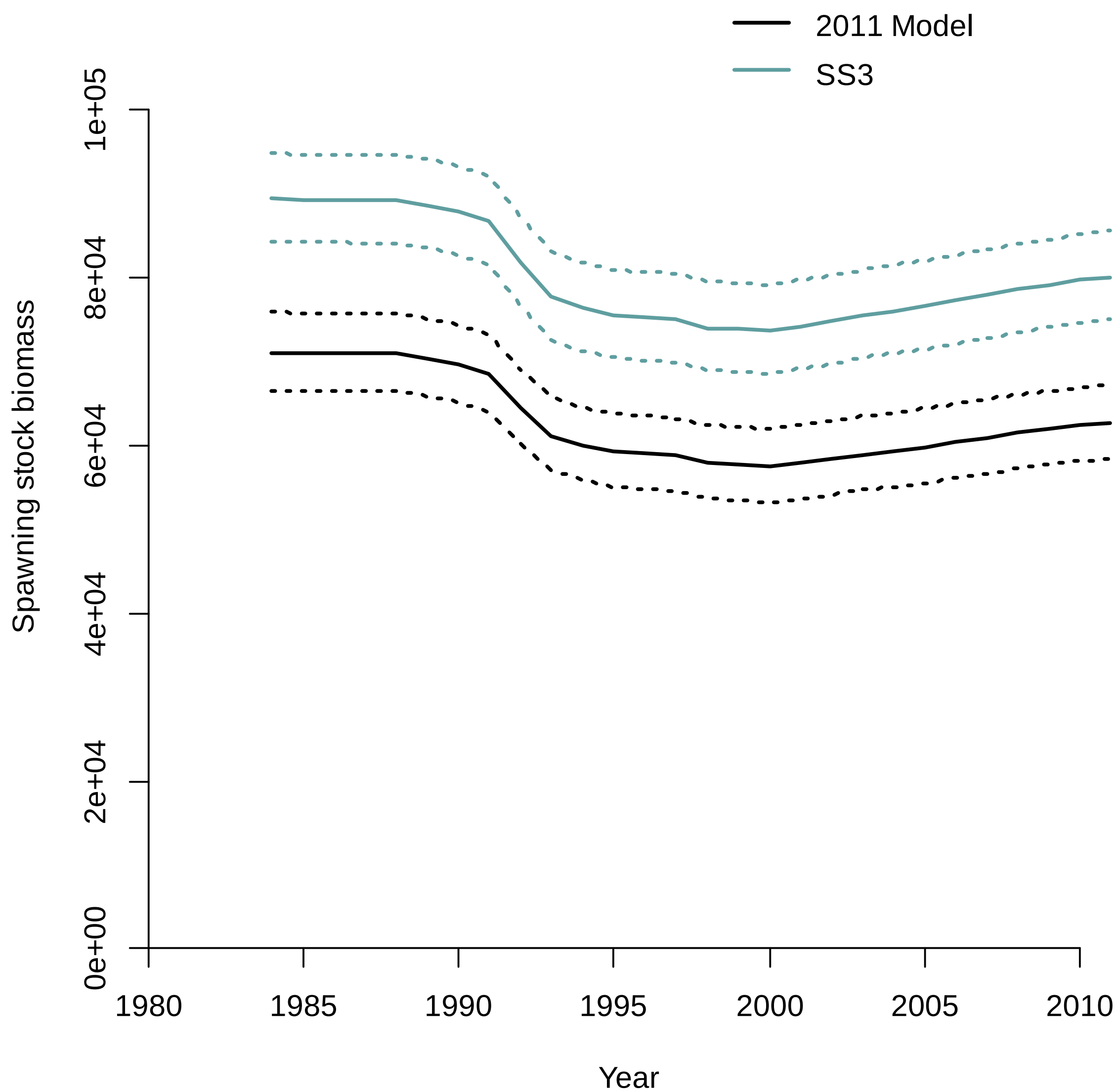


Figure 25. Spawning stock biomass for runs of the SS3 and 2011 models without estimation of recruitment deviations.

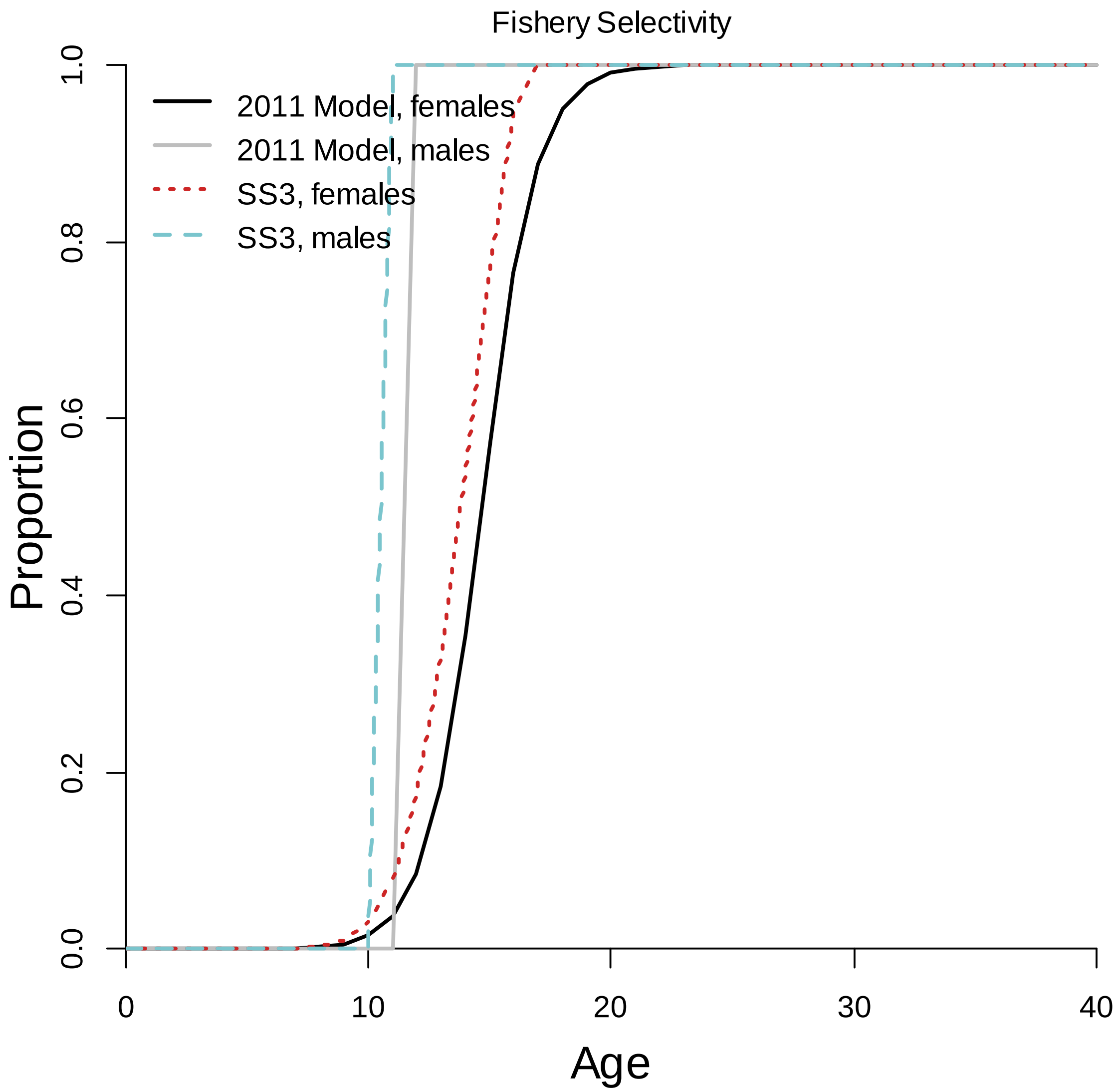


Figure 26. Fishery selectivity for runs of the SS3 and 2011 models without estimation of recruitment deviations.

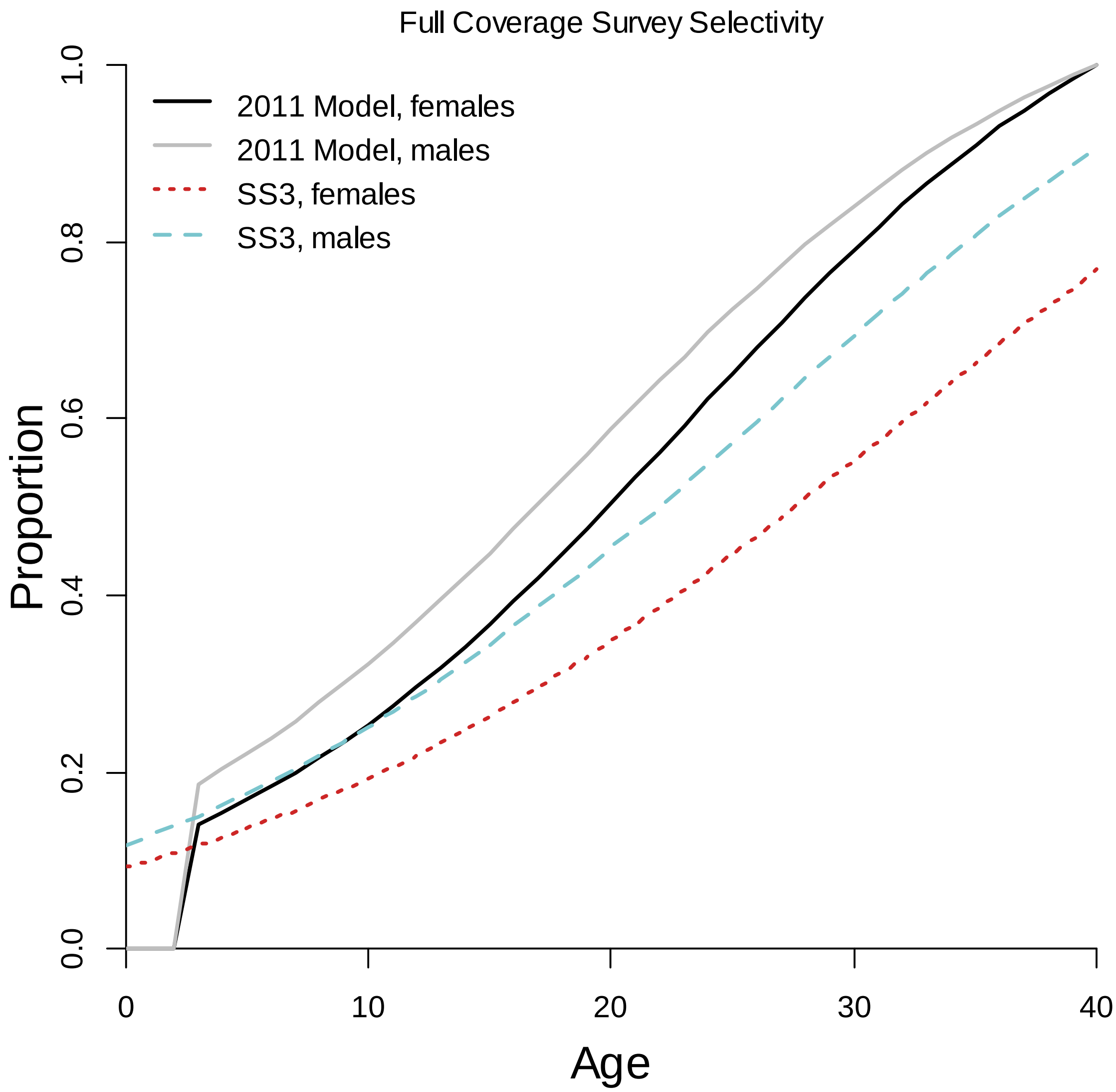


Figure 27. Full coverage survey selectivity for runs of the SS3 and 2011 models without estimation of recruitment deviations.

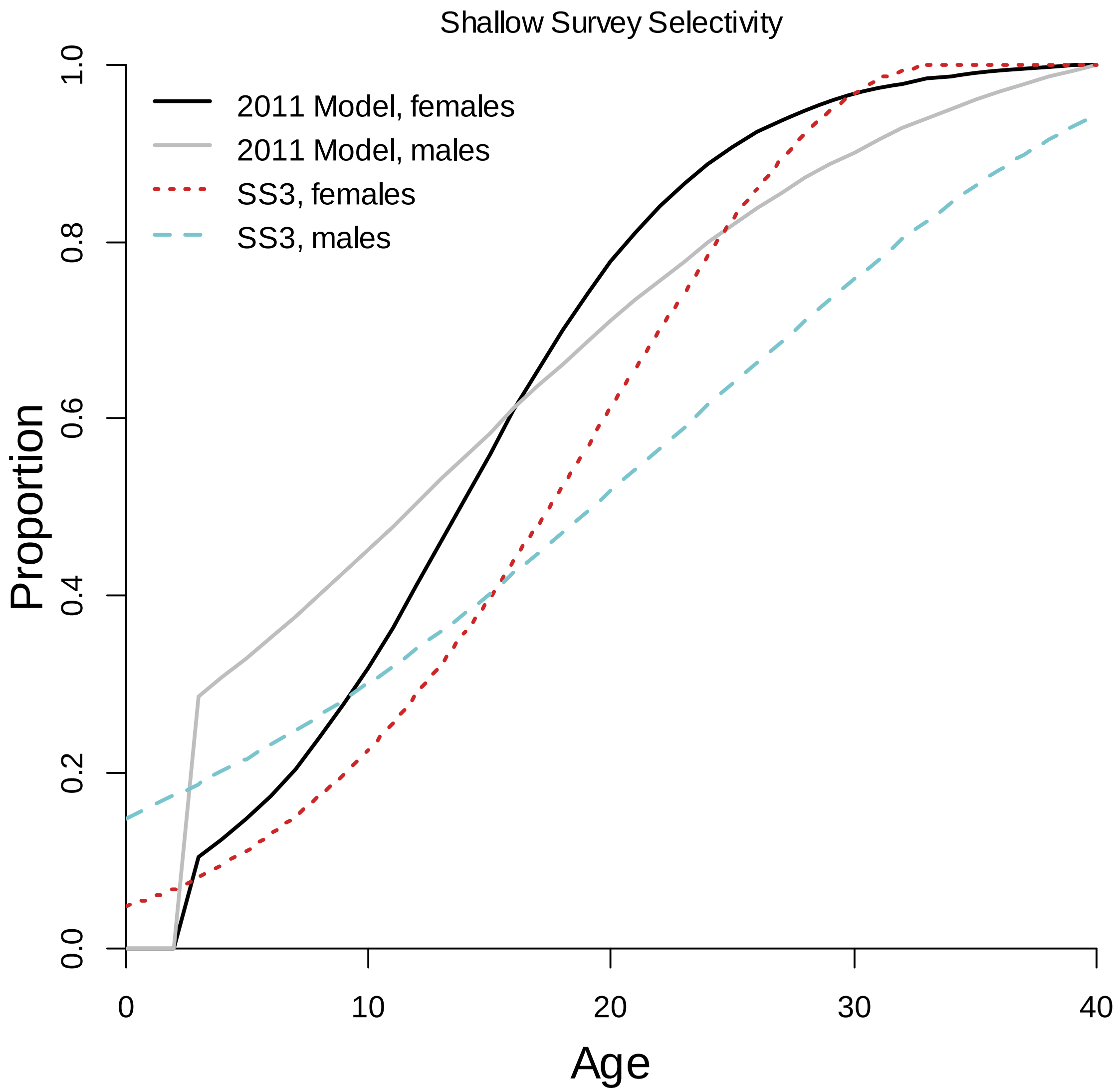


Figure 28. Shallow water survey selectivity for runs of the SS3 and 2011 models without estimation of recruitment deviations.

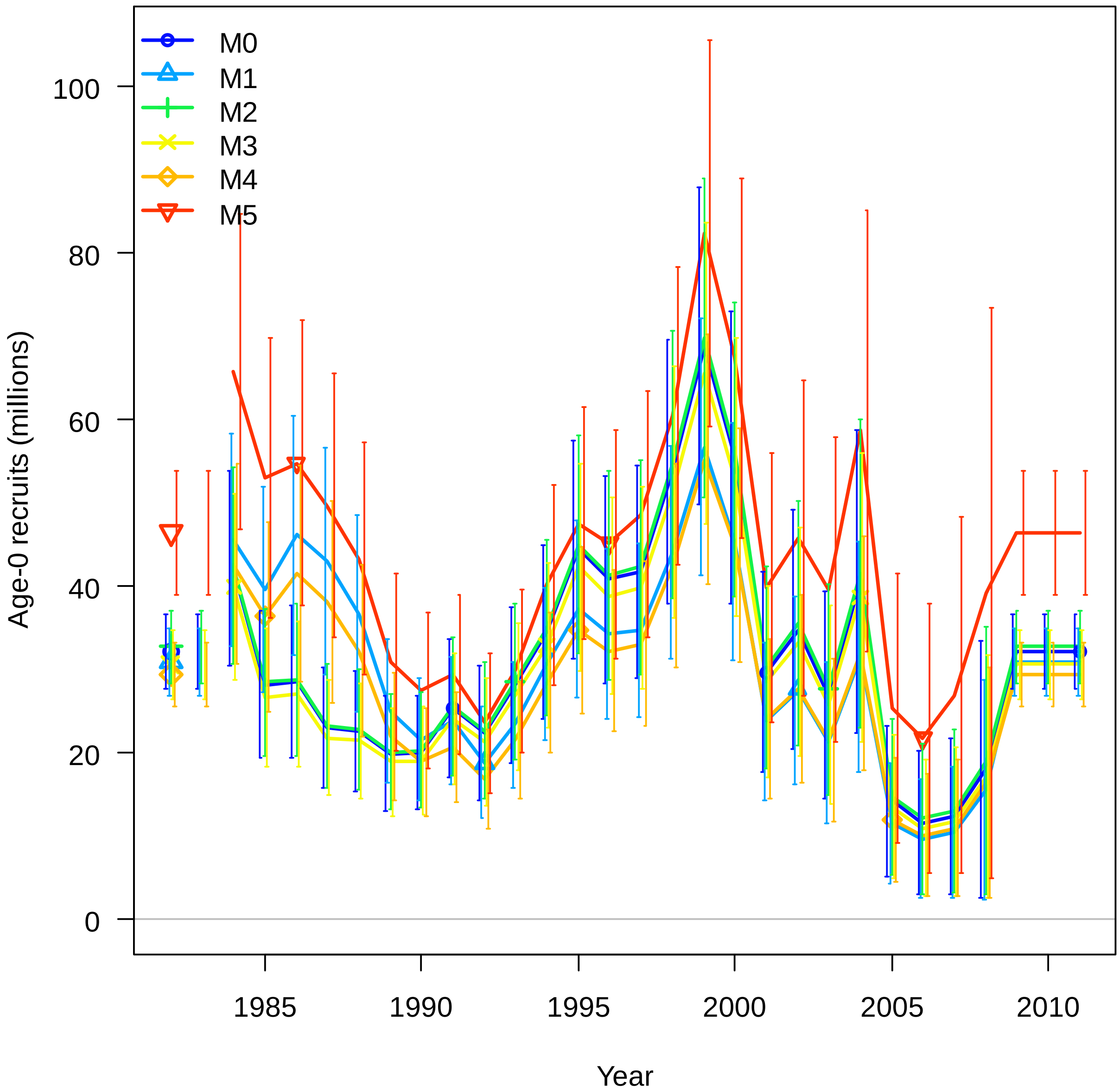


Figure 29. Age 0 recruits and 95% asymptotic confidence intervals for each alternative SS3 model. M0 is the transitional SS3 model that best matches the 2011 model. The leftmost group of vertical lines shows the log of mean recruitment.

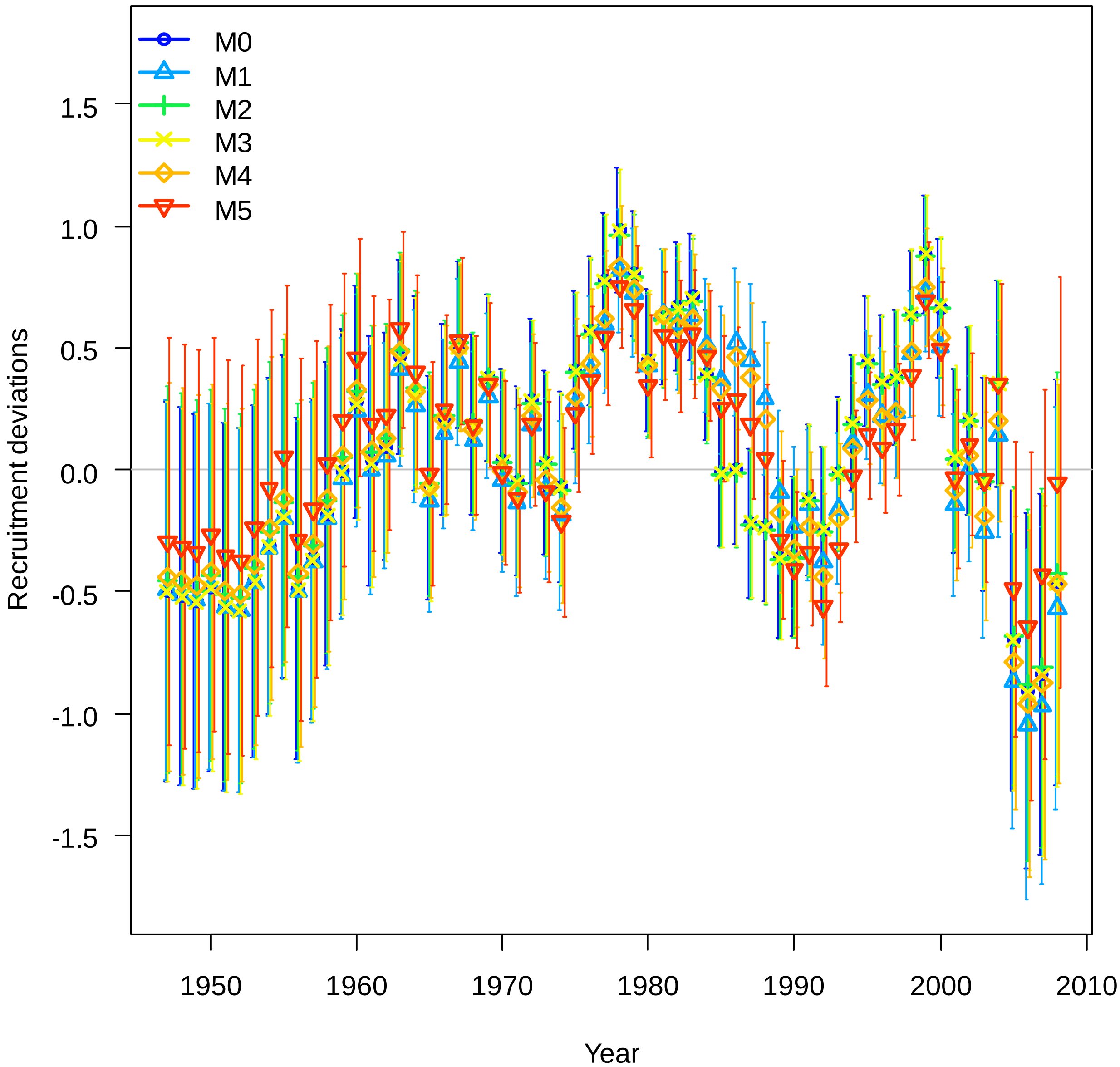


Figure 30. Estimated recruitment deviations and 95% asymptotic confidence intervals for each alternative SS3 model. M0 is the transitional SS3 model that best matches the 2011 model.

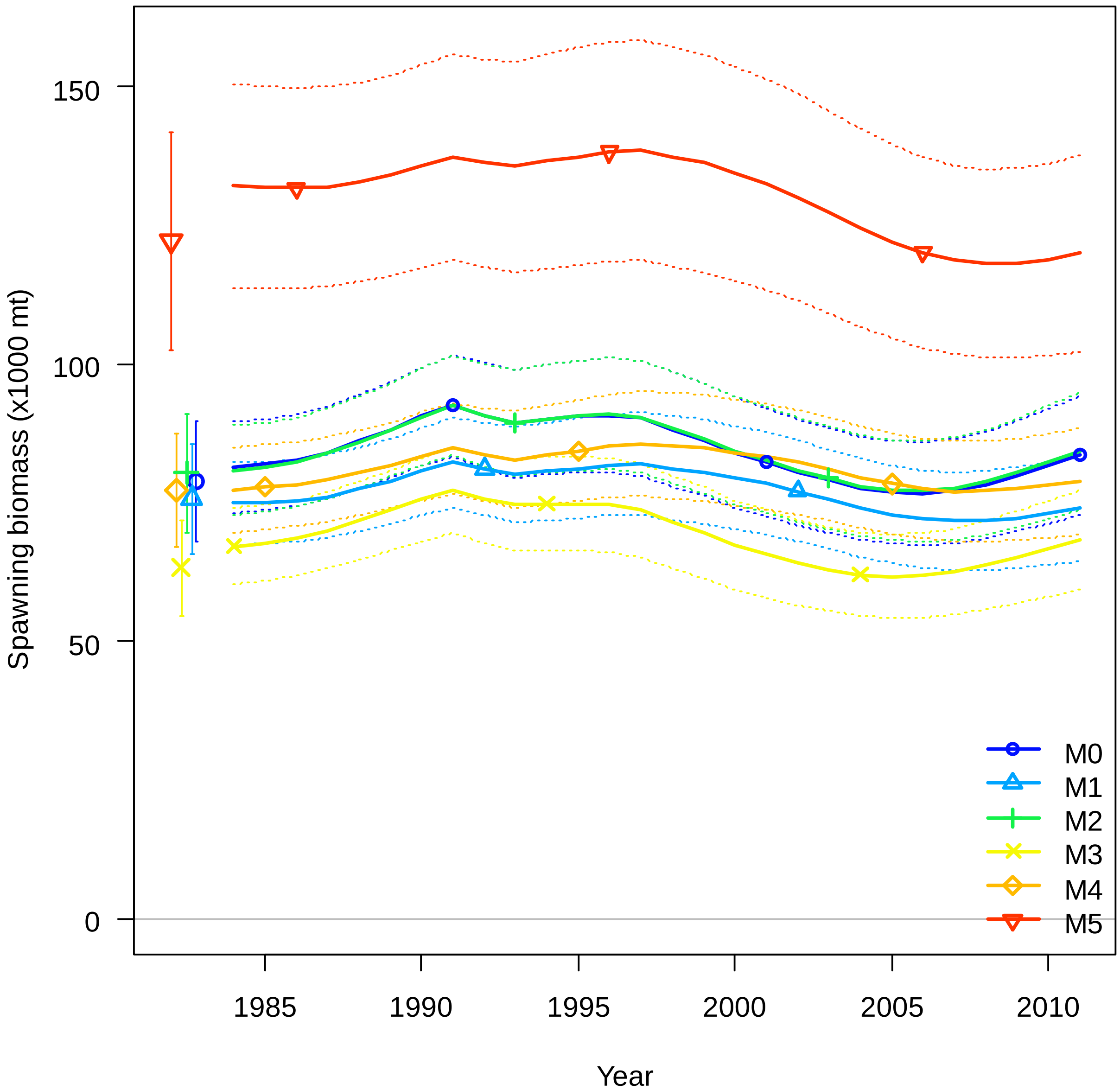


Figure 31. Spawning stock biomass (solid lines) and 95% asymptotic confidence intervals (dotted lines) over time for each alternative SS3 model. M0 is the transitional SS3 model that best matches the 2011 model.

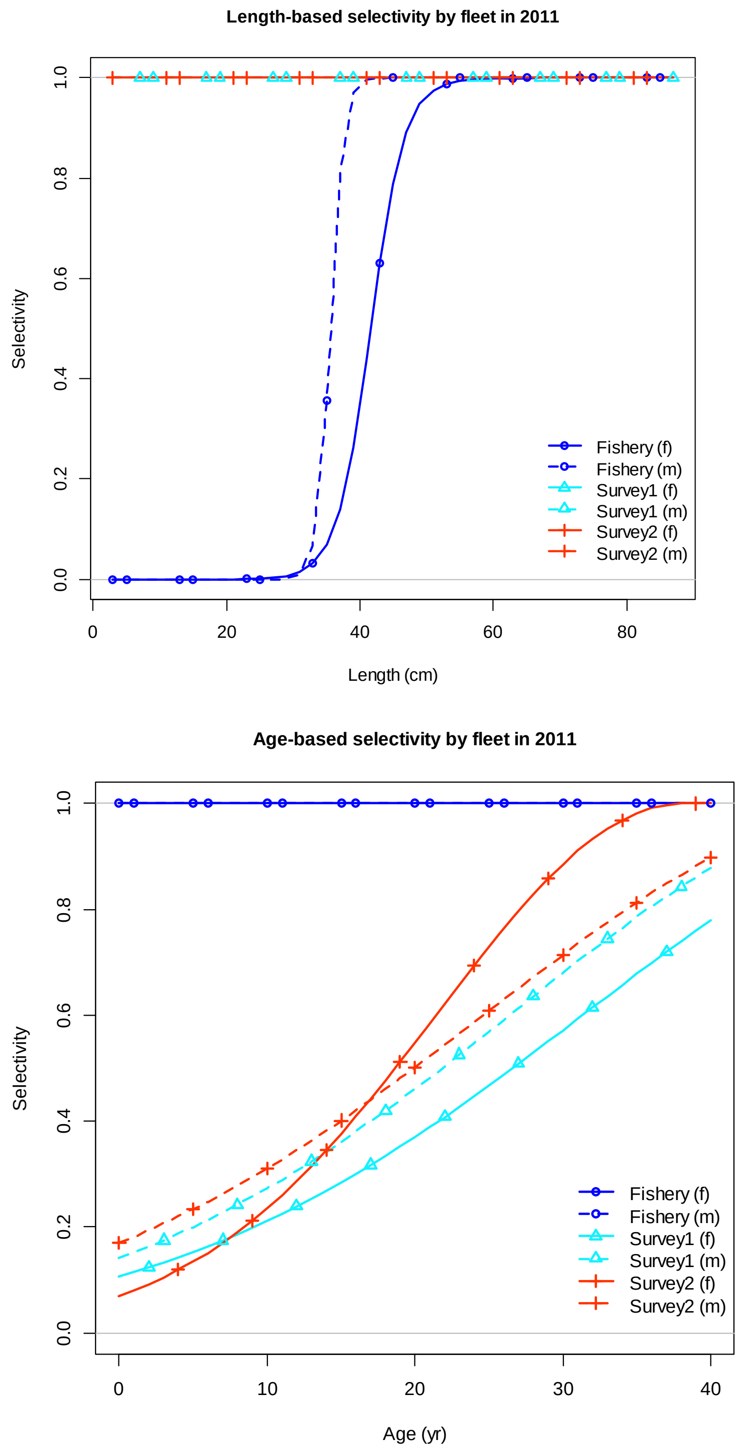


Figure 32. Length-based fishery selectivity (top panel) and age-based survey selectivity (bottom panel) for model M1 (as for the transitional SS3 model, but with length-based, logistic, sex-specific fishery selectivity).

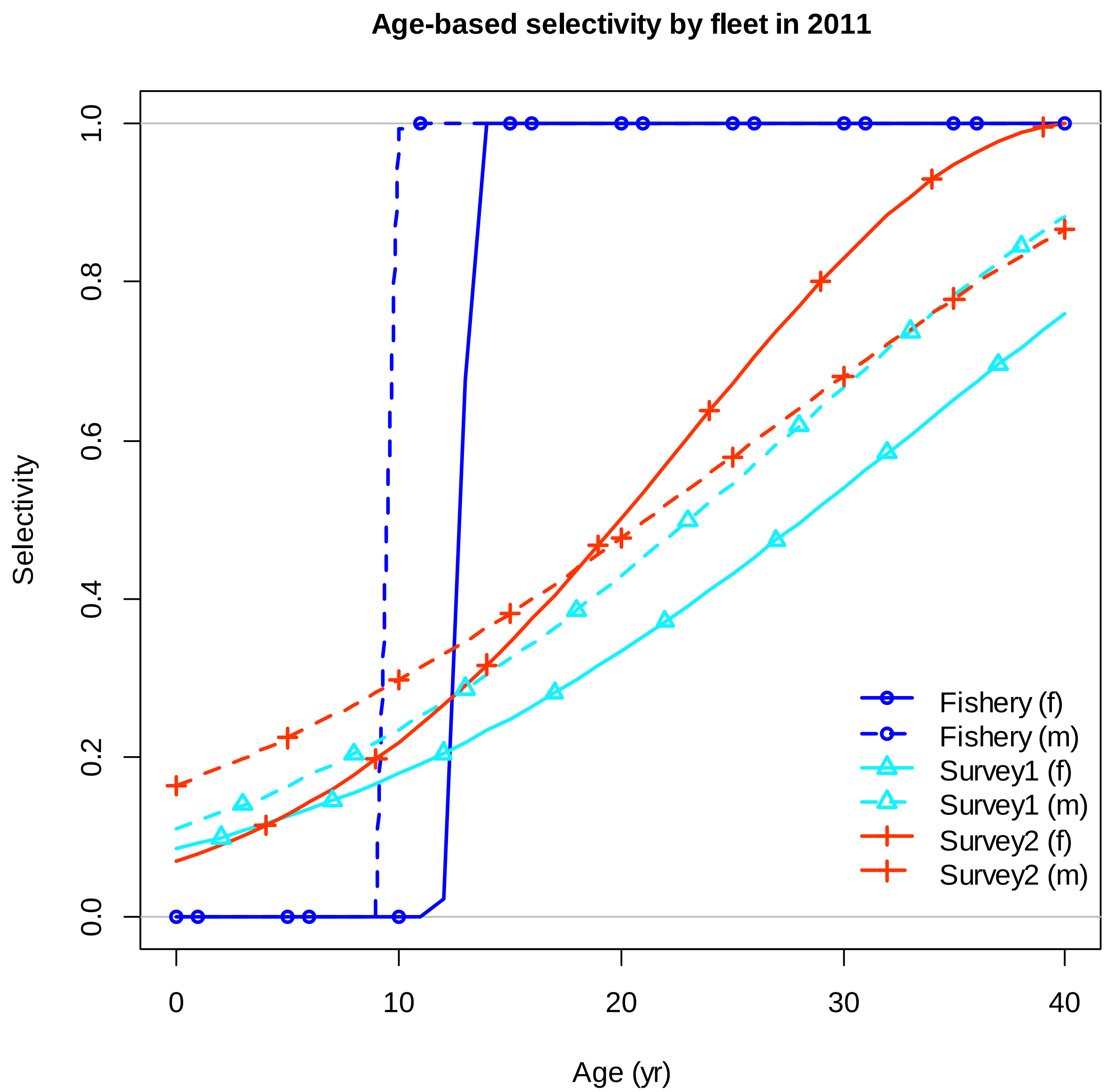


Figure 33. Fishery and survey selectivity curves for model M2 (as for the transitional SS3 model, but estimates an initial fishing mortality rate).

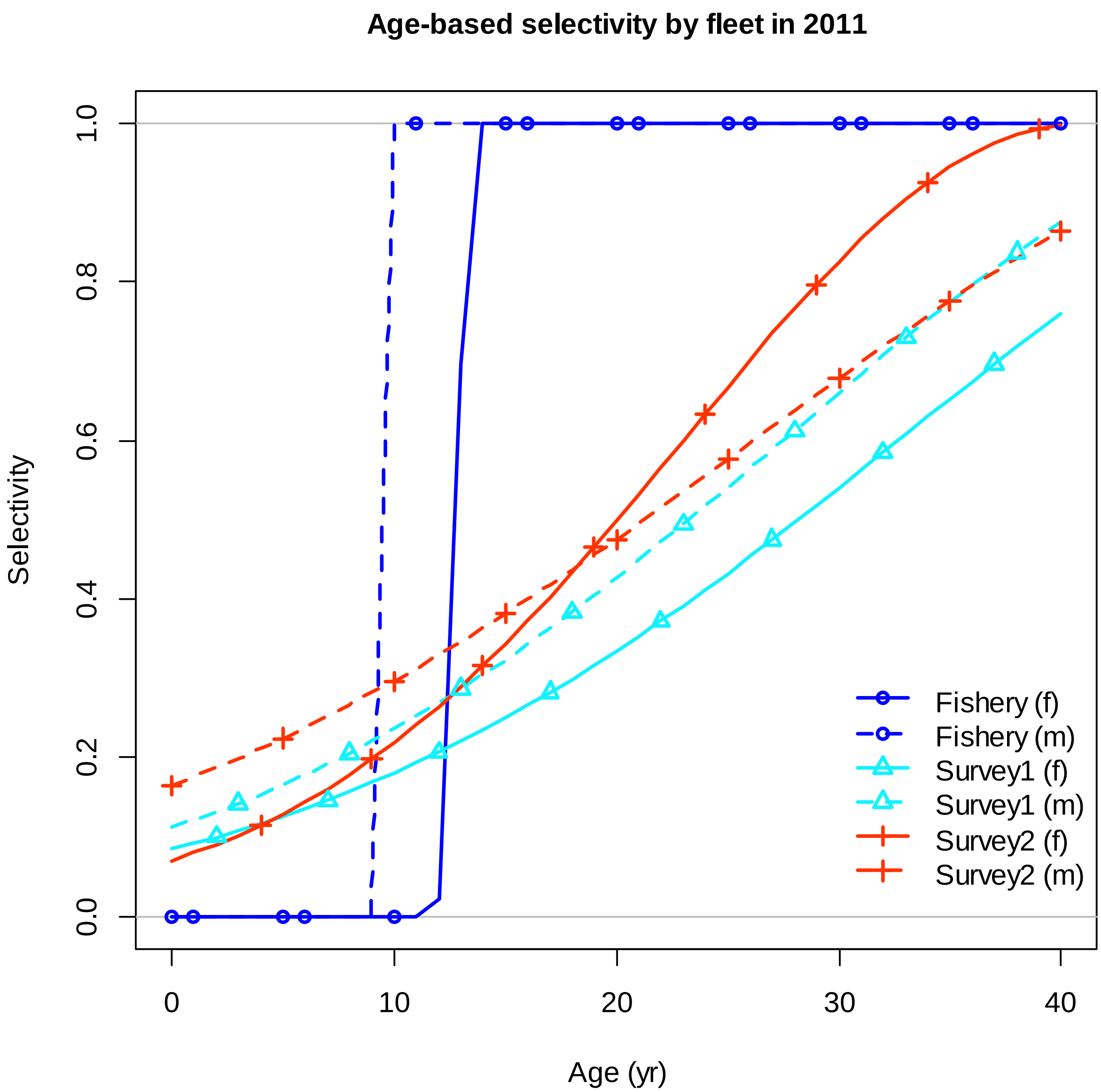


Figure 34. Selectivity curves for model M3 (as for the transitional SS3 model, but with fixed internal growth parameters specified).

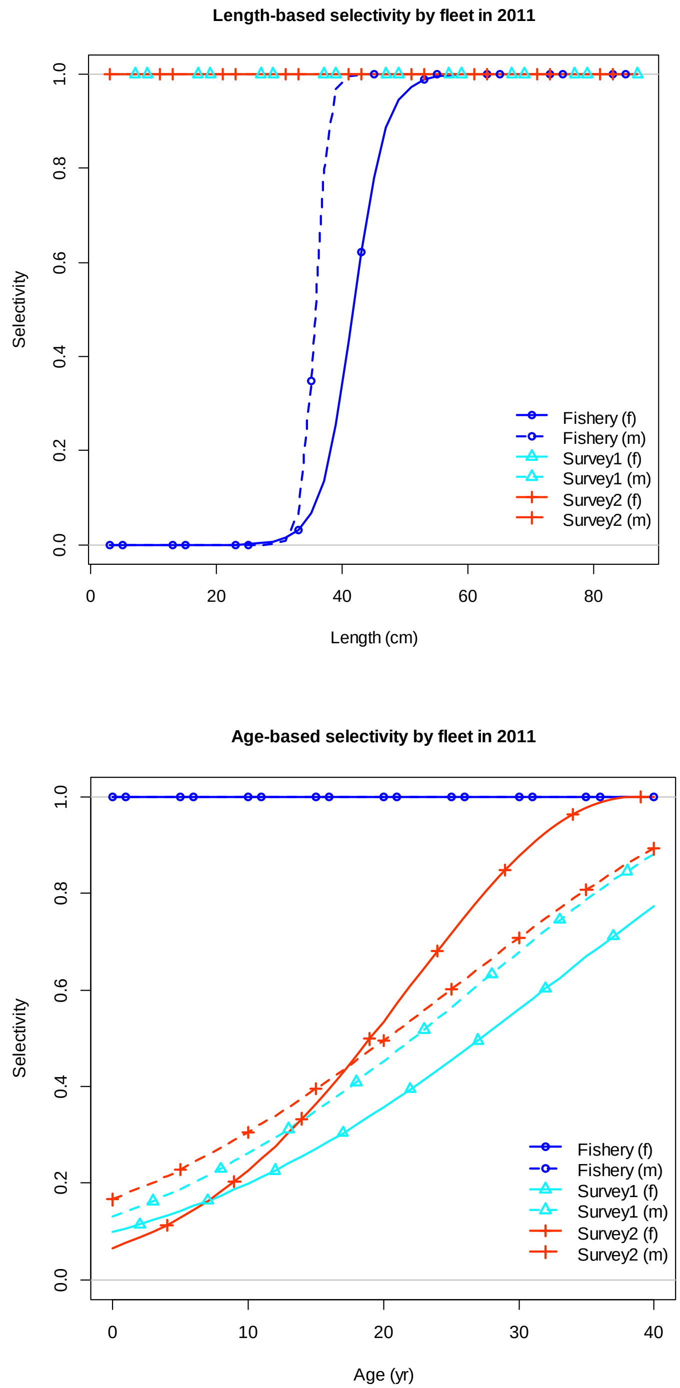


Figure 35. Selectivity curves for model M4 (as for the transitional SS3 model, but with fixed internal growth parameters, estimated initial equilibrium F, and length-based, logistic, sex-specific fishery selectivity). The top panel shows length-based fishery selectivity and the bottom panel shows age-based survey selectivity.

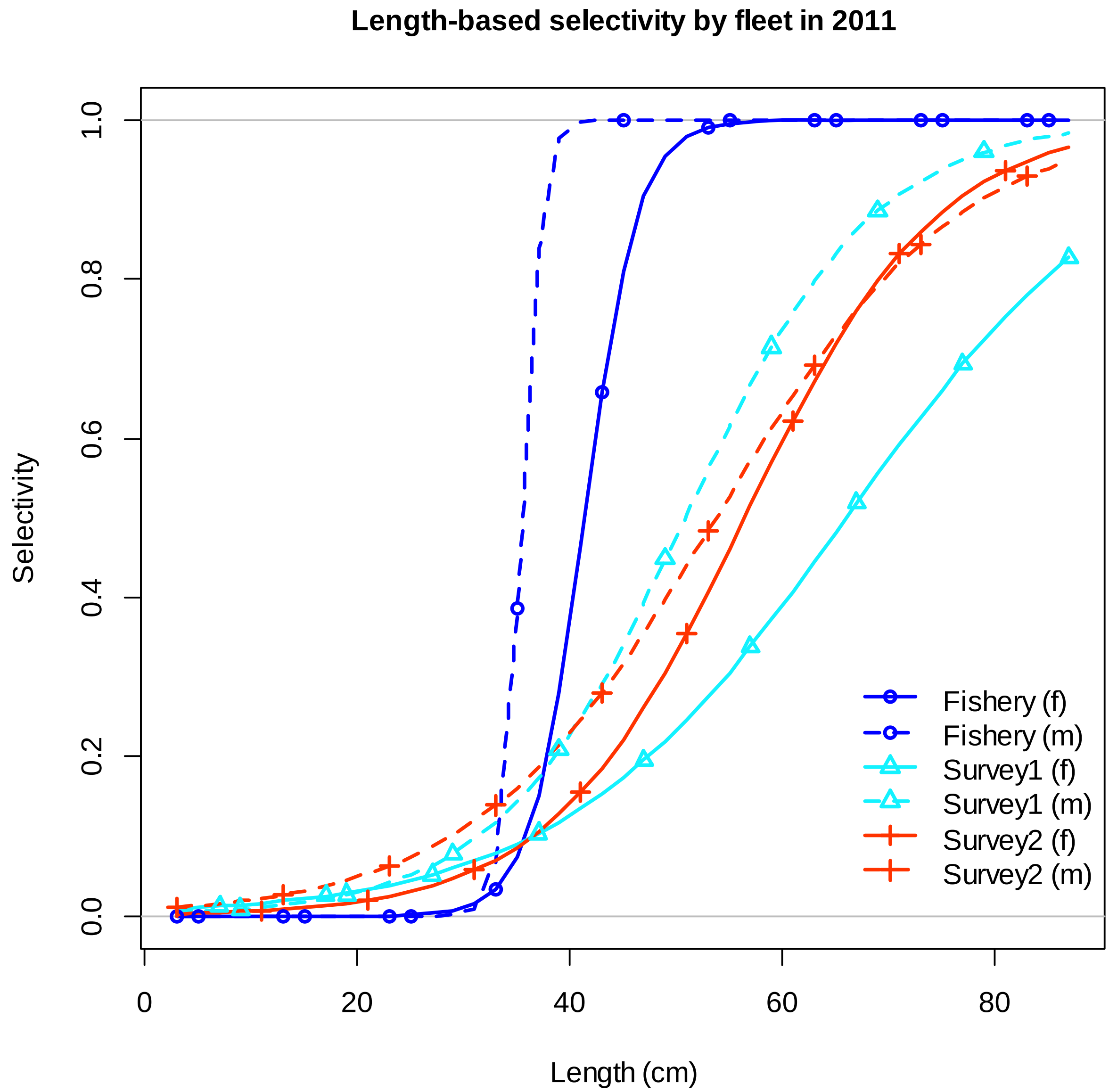


Figure 36. Length-based, logistic, sex-specific selectivity for the fishery, the full coverage survey (Survey1) and the shallow-water survey (Survey2) for model M5 (M5: internal, fixed growth parameters, estimation of initial equilibrium F, and length-based selectivity for the fishery and both surveys).

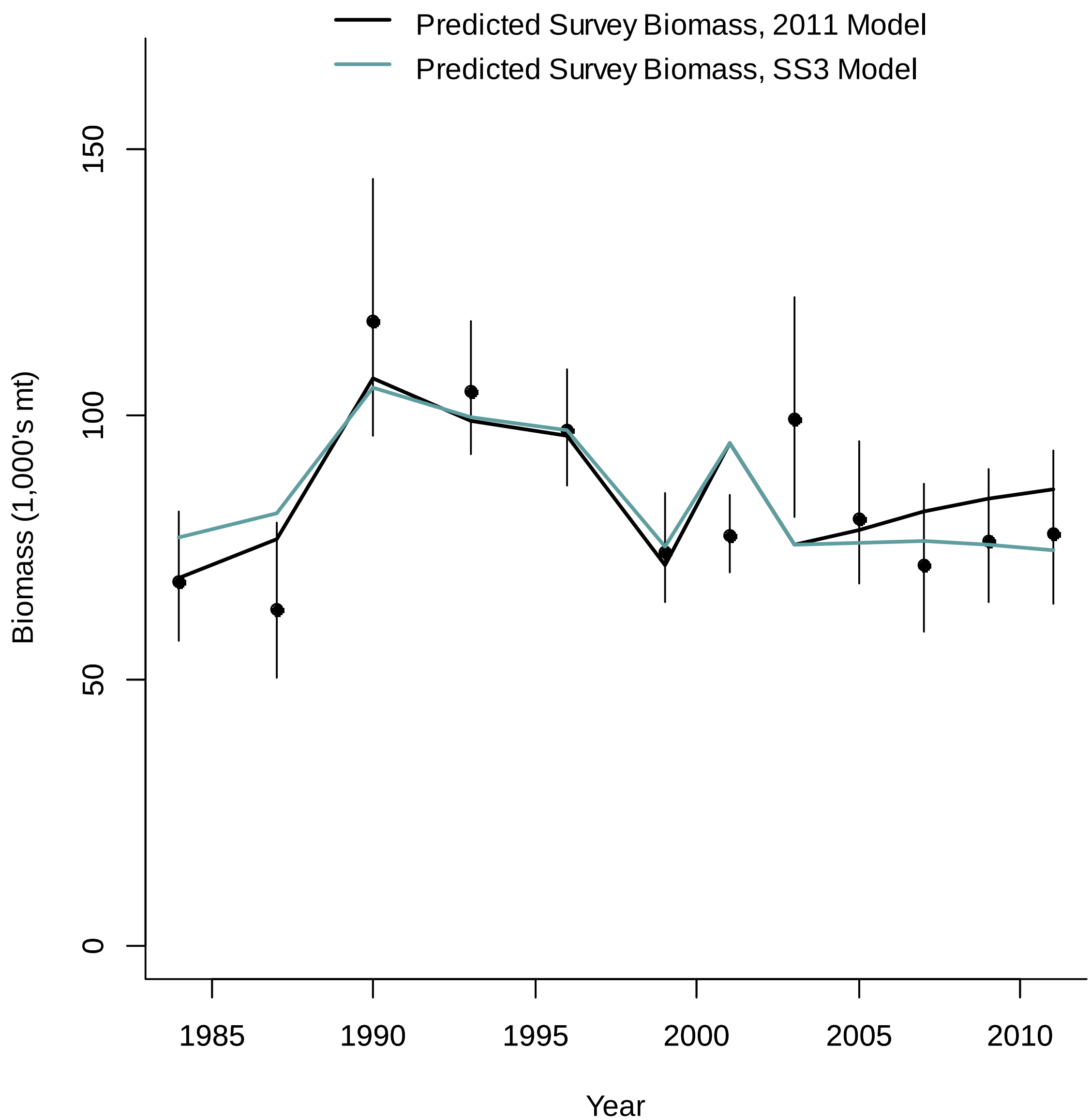
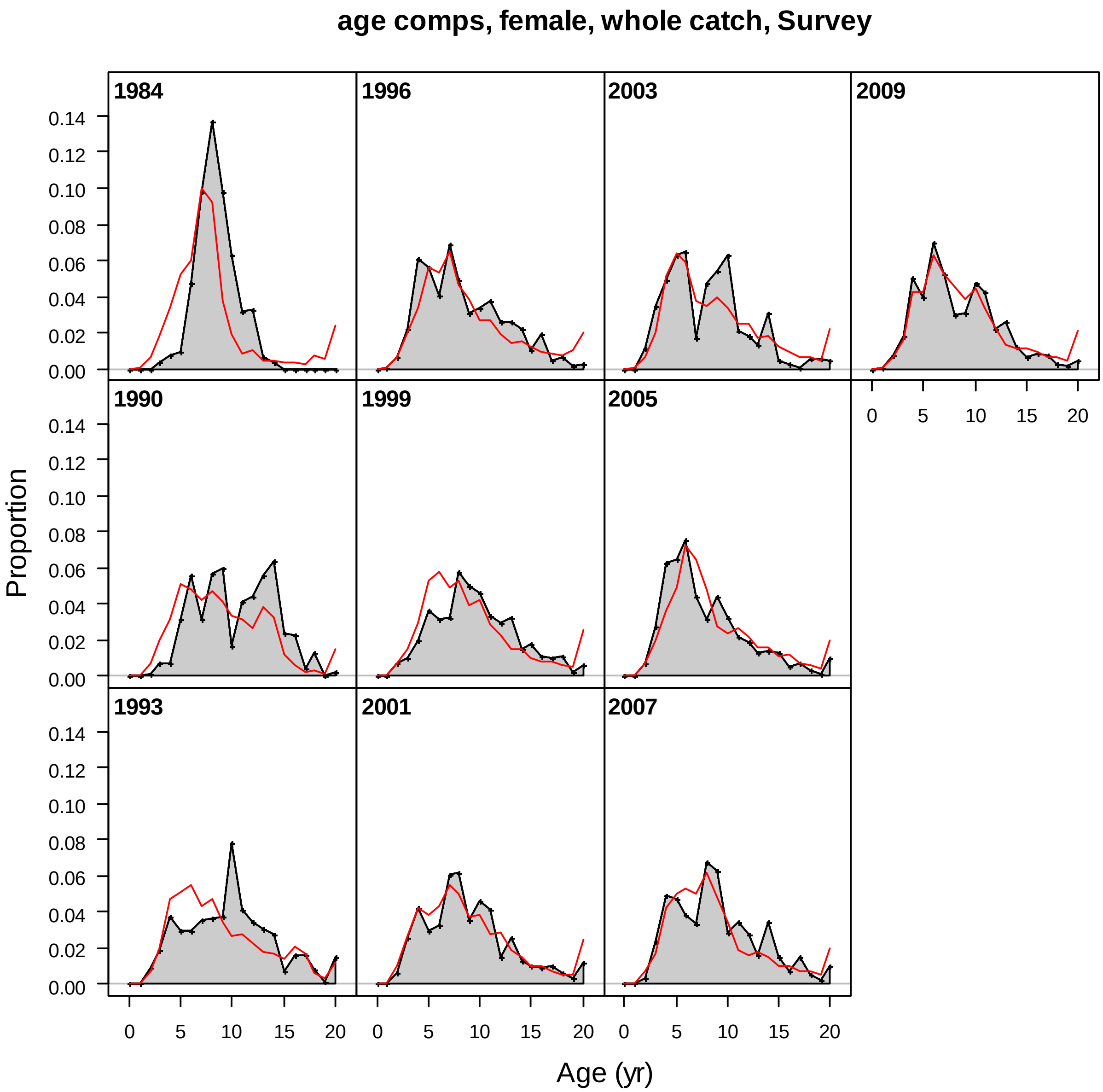


Figure 37. Observed survey biomass (black dots) with 95% asymptotic confidence intervals (vertical black lines) and predicted survey biomass for the proposed alternative model M4 (as for the transitional SS3 model, but with fixed internal growth parameters, estimated initial equilibrium F, and length-based, logistic, sex-specific fishery selectivity; blue lines) and the 2011 model (black lines).


age comps, female, whole catch, Survey
1984
1996
2003
2009
1990
1999
2005
1993
2001
2007
Proportion
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0
5
10
15
20
Age (yr)

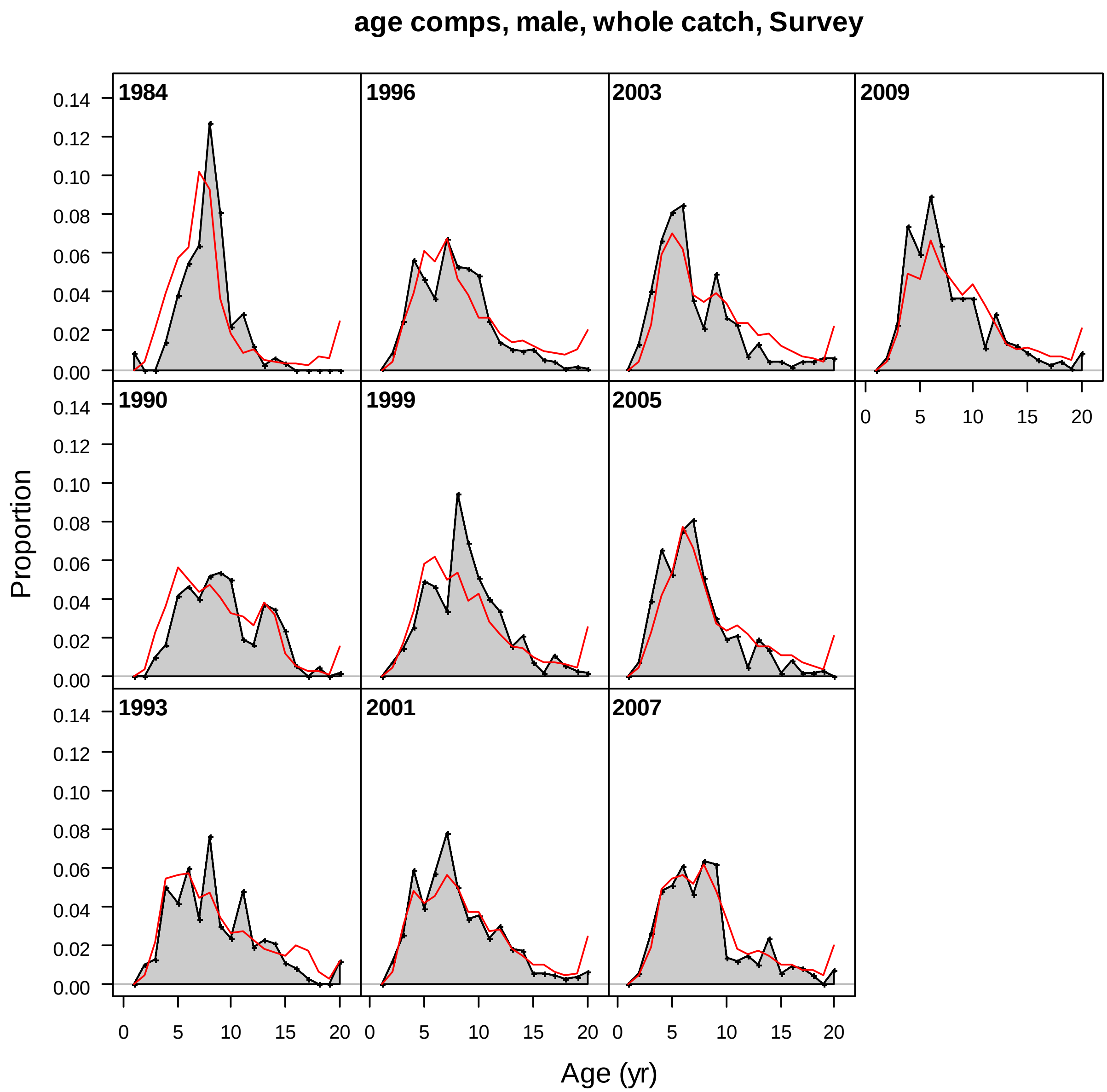
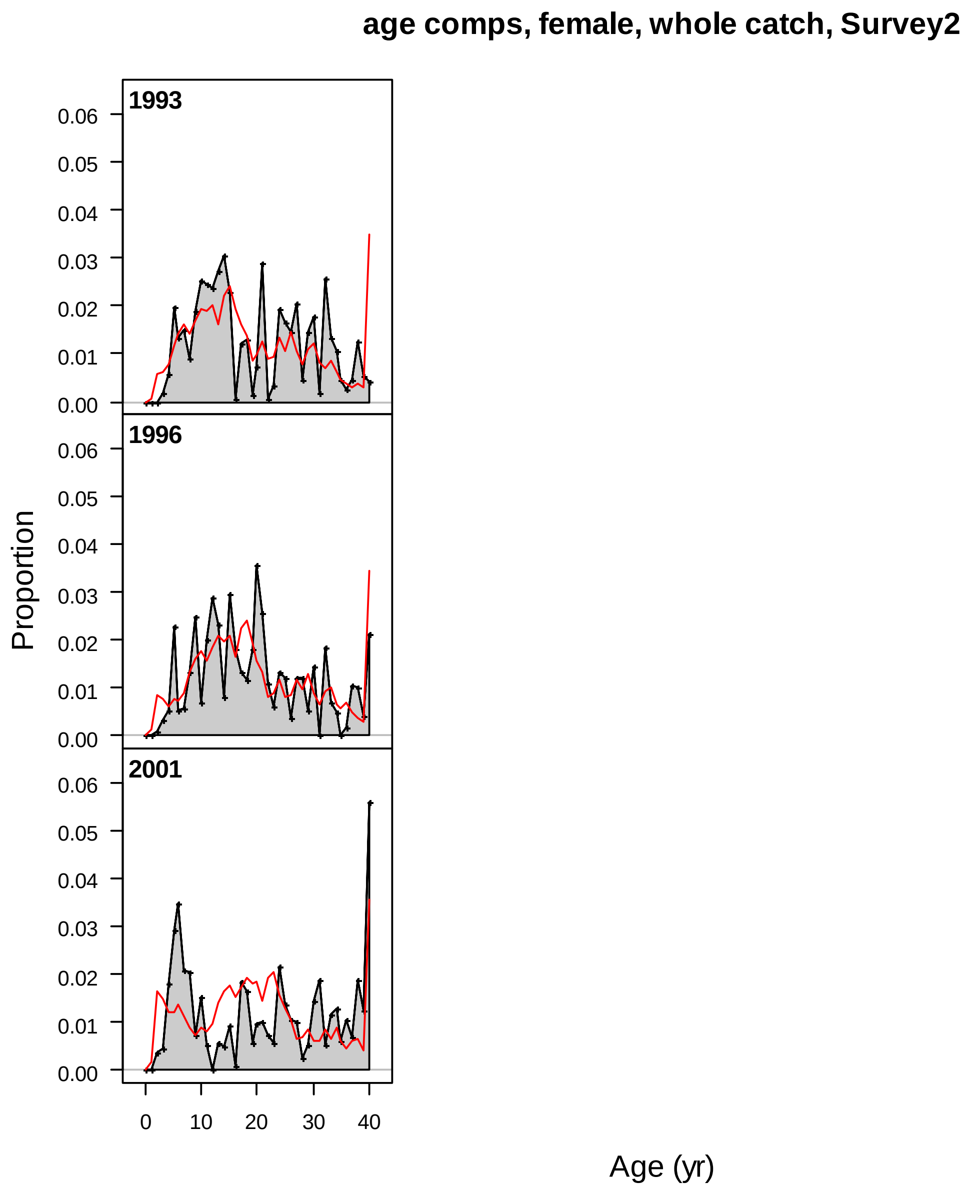


Figure 38. Observed (solid black lines and grey shaded area) and predicted (red lines) full-coverage survey proportions-at-age for proposed alternative model M4 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery selectivity is a logistic, sex-specific, length-based function) for females (first panel) and males (second panel).


age comps, female, whole catch, Survey2
1993
1996
2001
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Proportion
0
10
20
30
40
Age (yr)

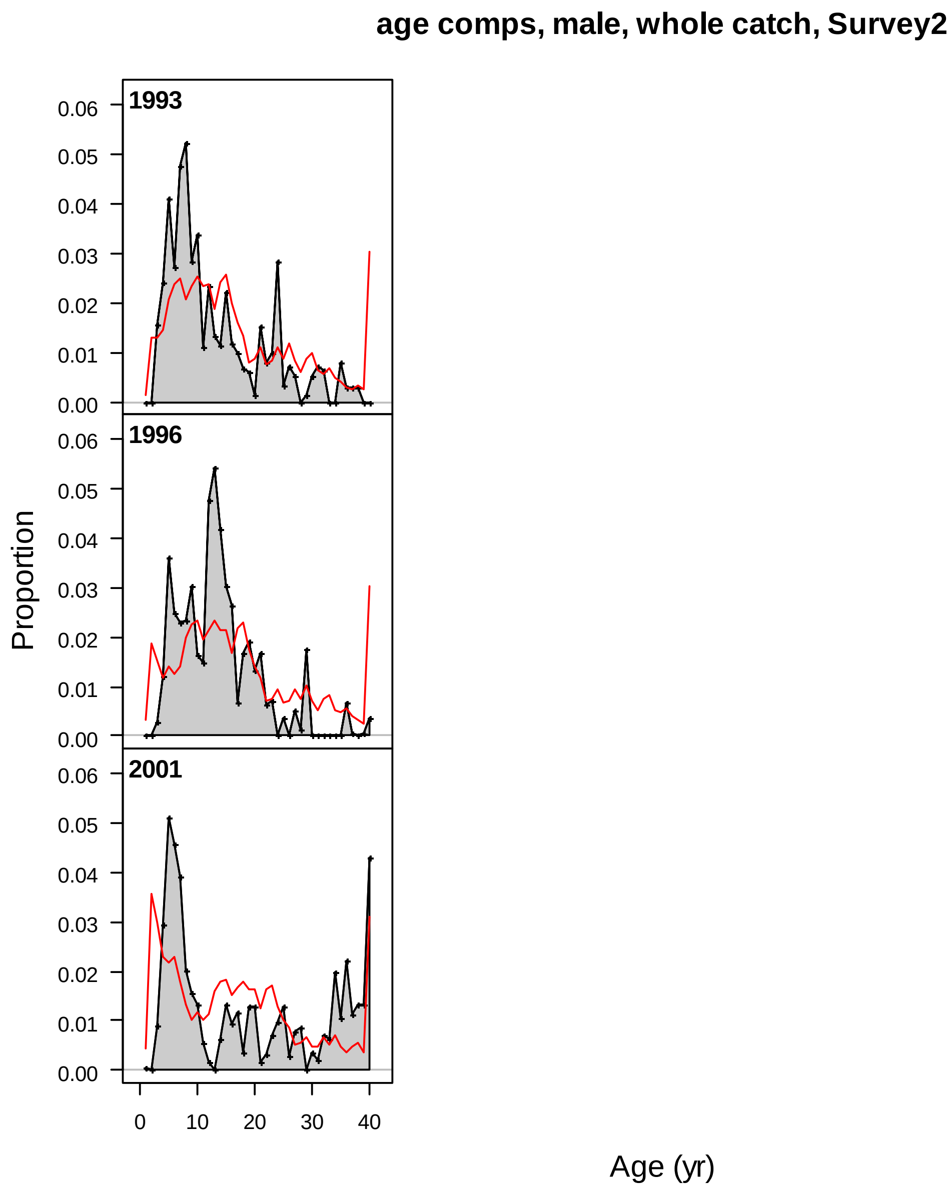
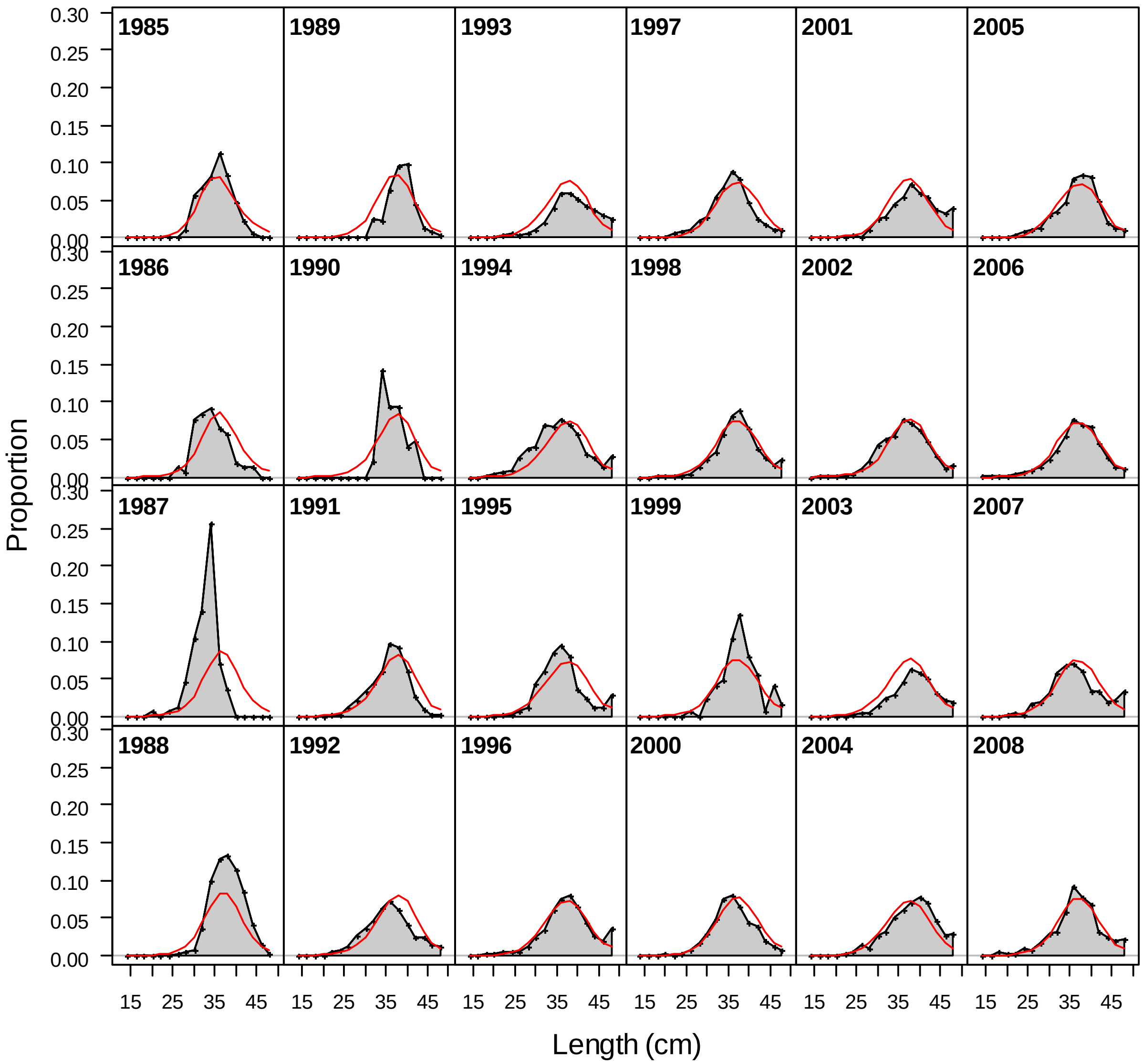
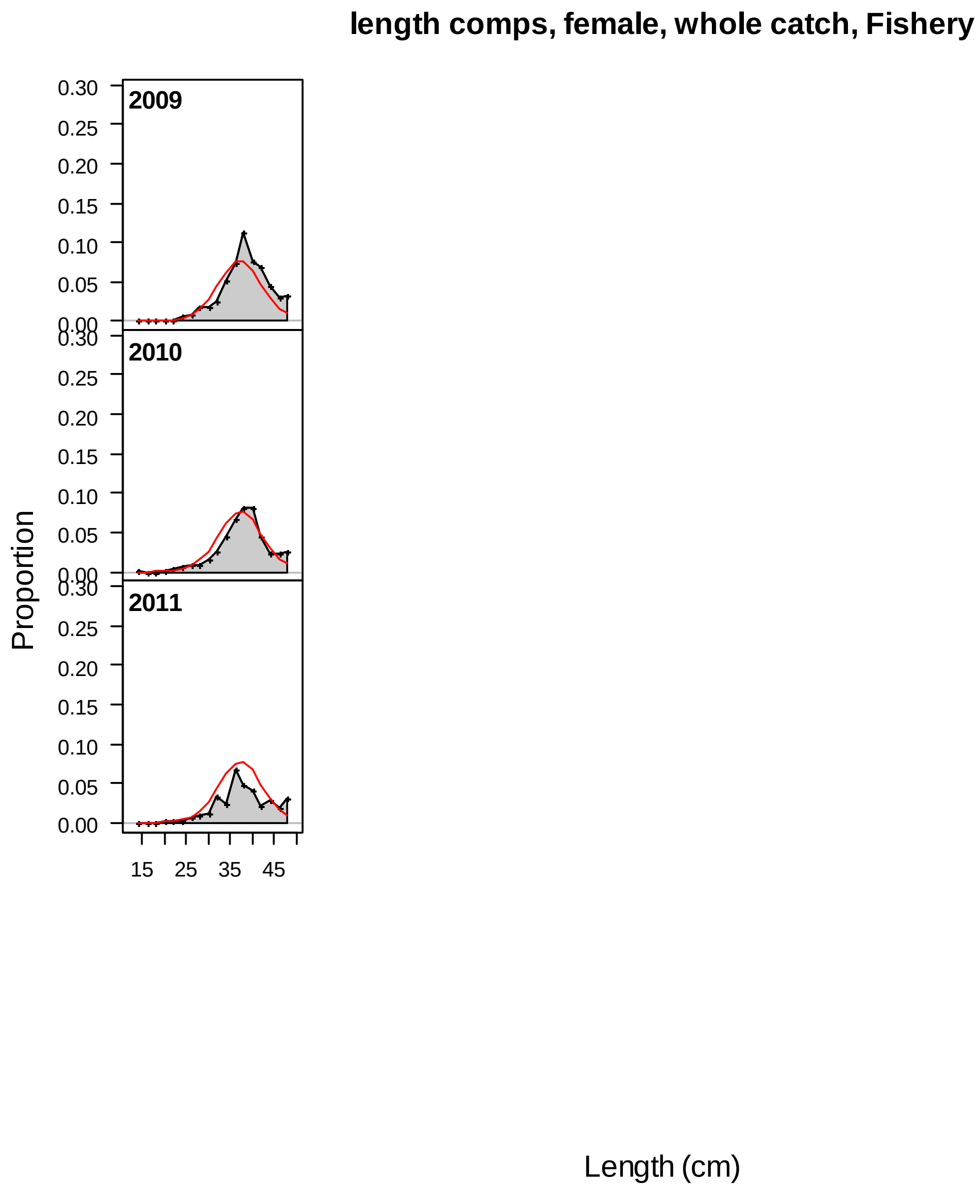
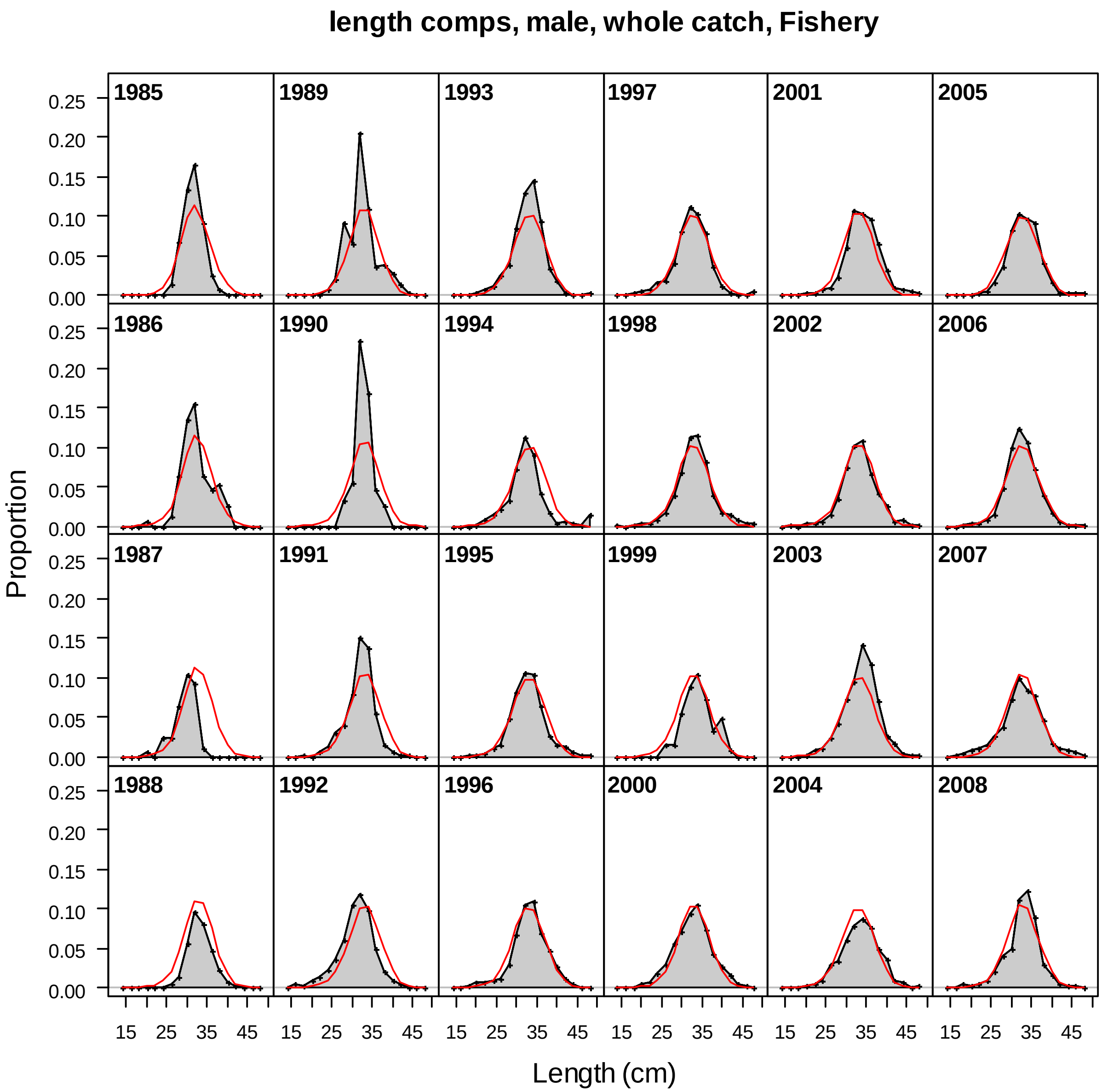


Figure 39. Observed (solid black lines and grey shaded area) and predicted (red lines) shallow-water survey proportions-at-age for proposed alternative model M4 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery selectivity is a logistic, sex-specific, length-based function) for females (first panel) and males (second panel).

length comps, female, whole catch, Fishery




length comps, female, whole catch, Fishery
2009
2010
2011
Proportion
0.30
0.25
0.20
0.15
0.10
0.05
0.00
15
25
35
45
Length (cm)


length comps, male, whole catch, Fishery
1985
1989
1993
1997
2001
2005
1986
1990
1994
1998
2002
2006
1987
1991
1995
1999
2003
2007
1988
1992
1996
2000
2004
2008
Proportion
0.00
0.05
0.10
0.15
0.20
0.25
15
25
35
45
Length (cm)

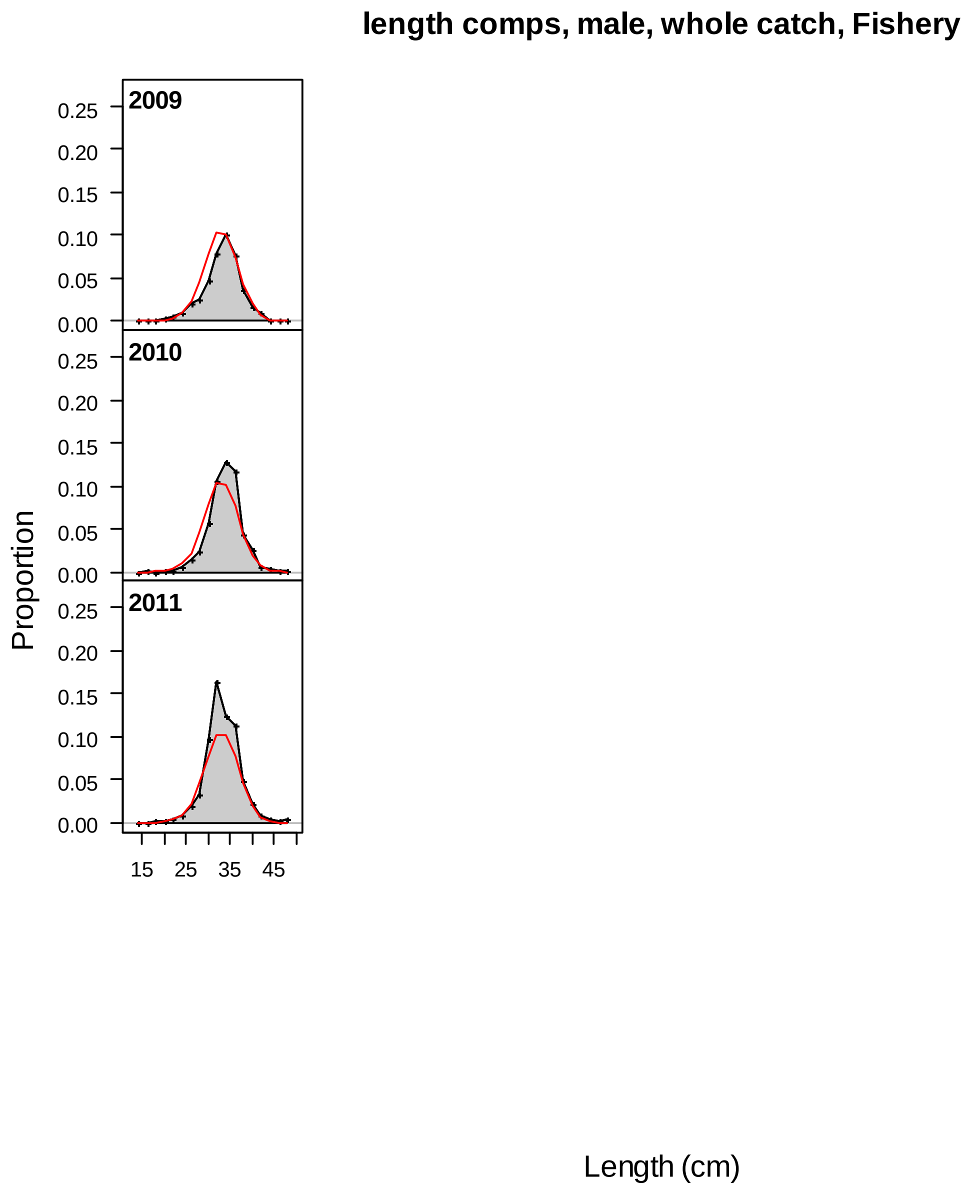
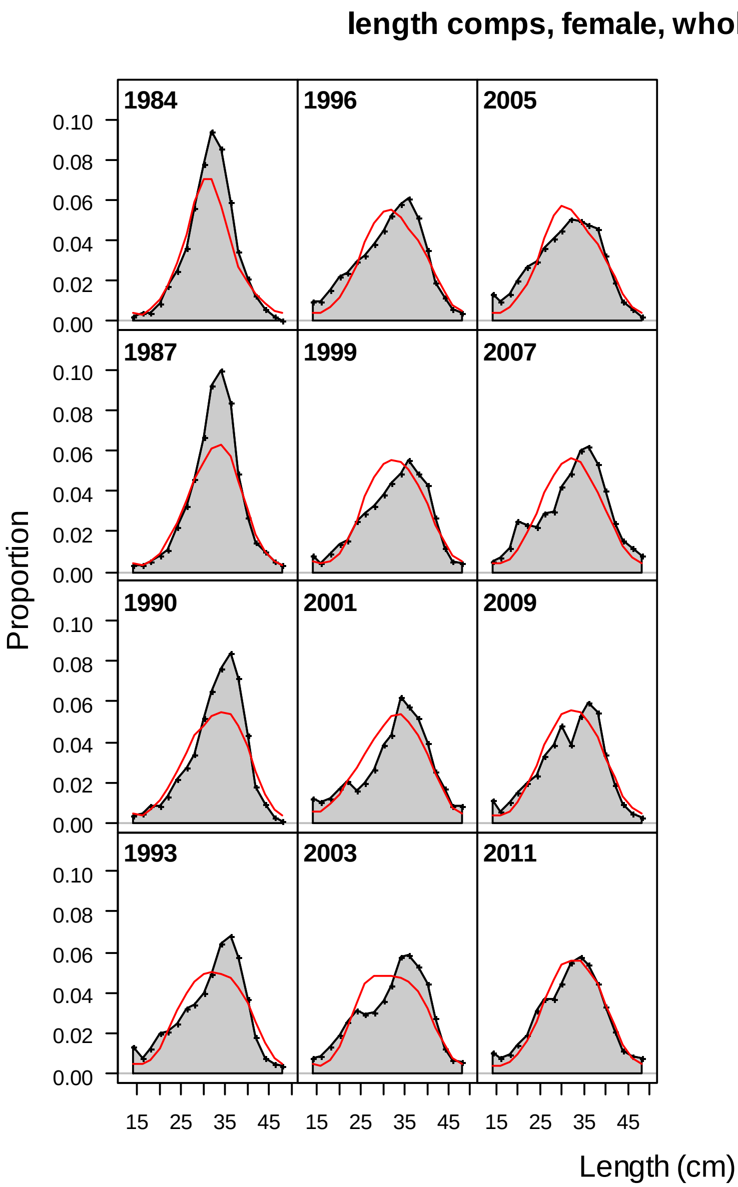


Figure 40. Observed (solid black lines and grey shaded area) and predicted (red lines) fishery proportions-at-length proposed alternative model M4 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery selectivity is a logistic, sex-specific, length-based function) for females (first set of panels) and males (second set of panels).


length comps, female, whole catch, Survey
1984
1996
2005
1987
1999
2007
1990
2001
2009
1993
2003
2011
Proportion
0.00
0.02
0.04
0.06
0.08
0.10
15
25
35
45
Length (cm)

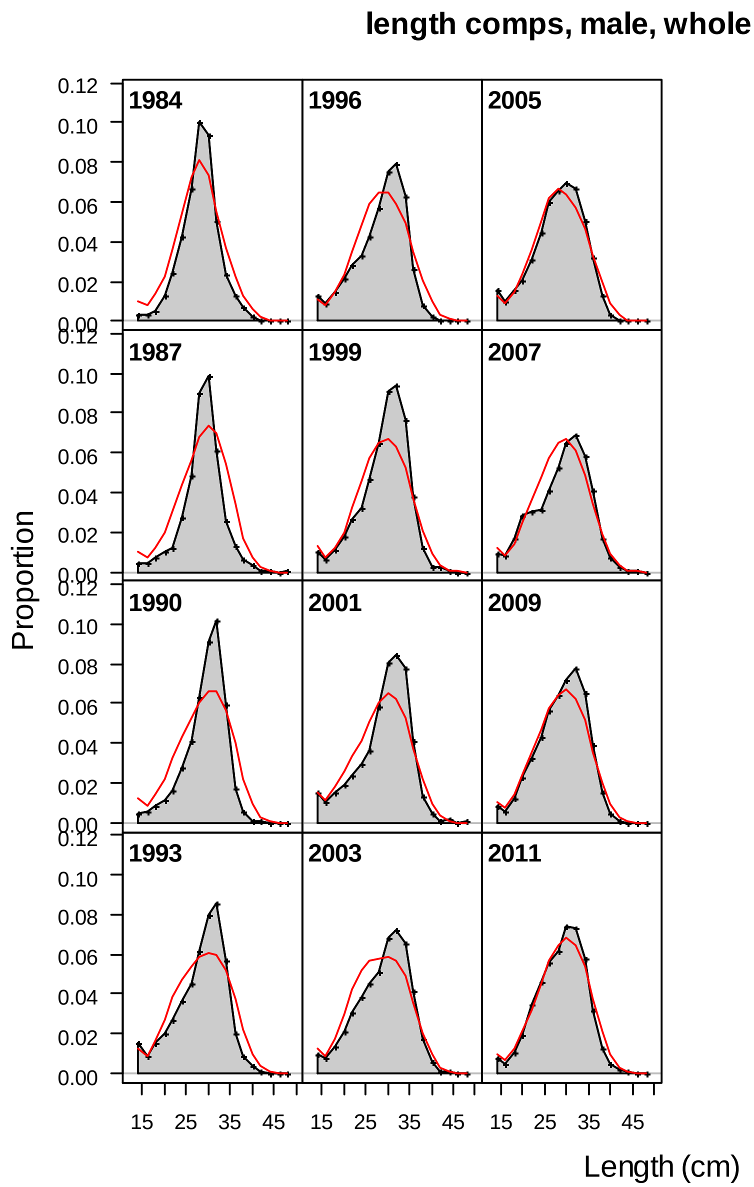
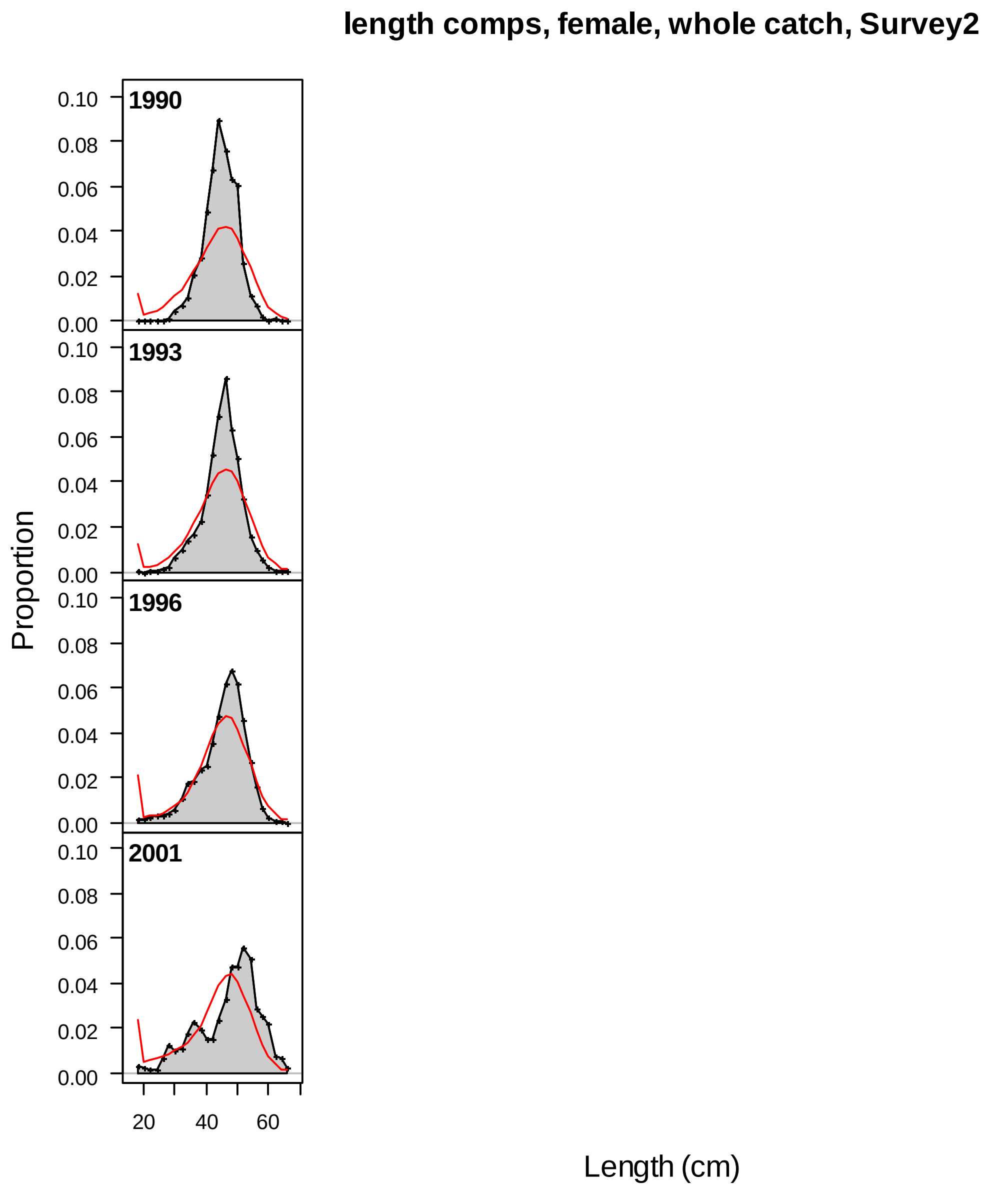


Figure 41. Observed (solid black lines and grey shaded area) and predicted (red lines) full-coverage survey proportions-at-length for proposed alternative model M4 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery selectivity is a logistic, sex-specific, length-based function) for females (first set of panels) and males (second set of panels).


length comps, female, whole catch, Survey2
1990
1993
1996
2001
0.10
0.08
0.06
0.04
0.02
0.00
Proportion
20
40
60
Length (cm)

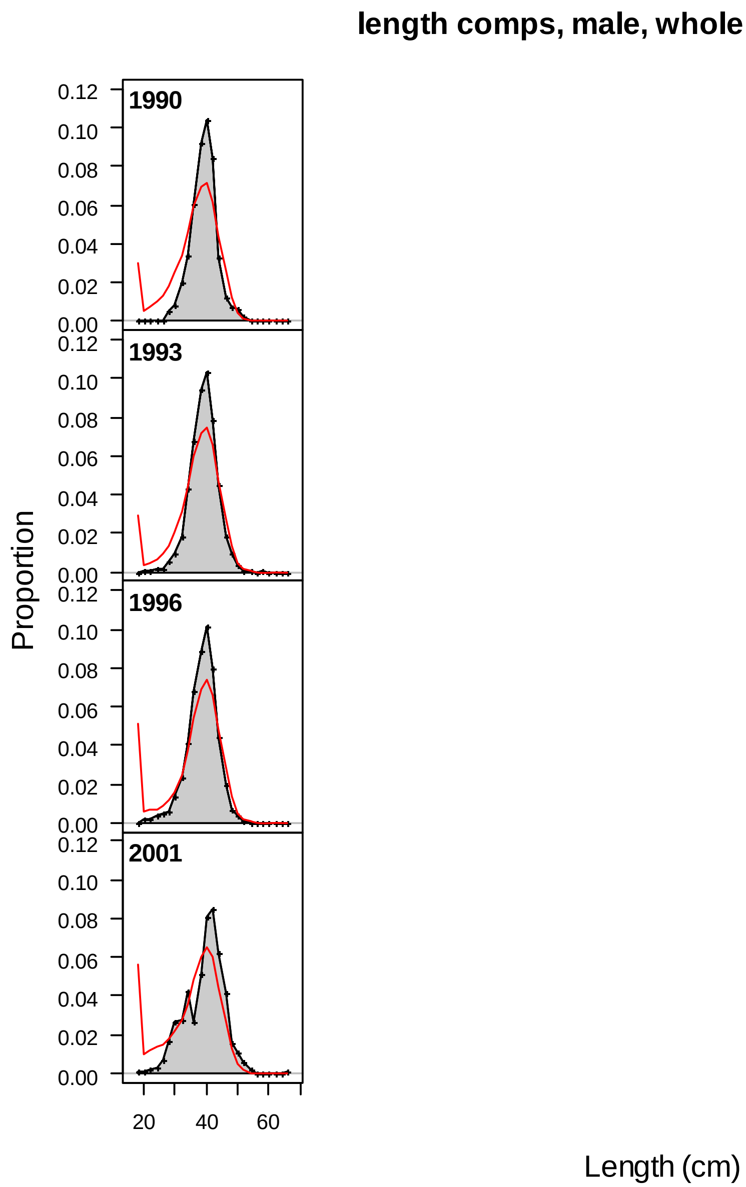


Figure 42. Observed (solid black lines and grey shaded area) and predicted (red lines) shallow-water survey proportions-at-length for proposed alternative model M4 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery selectivity is a logistic, sex-specific, length-based function) for females (first set of panels) and males (second set of panels).

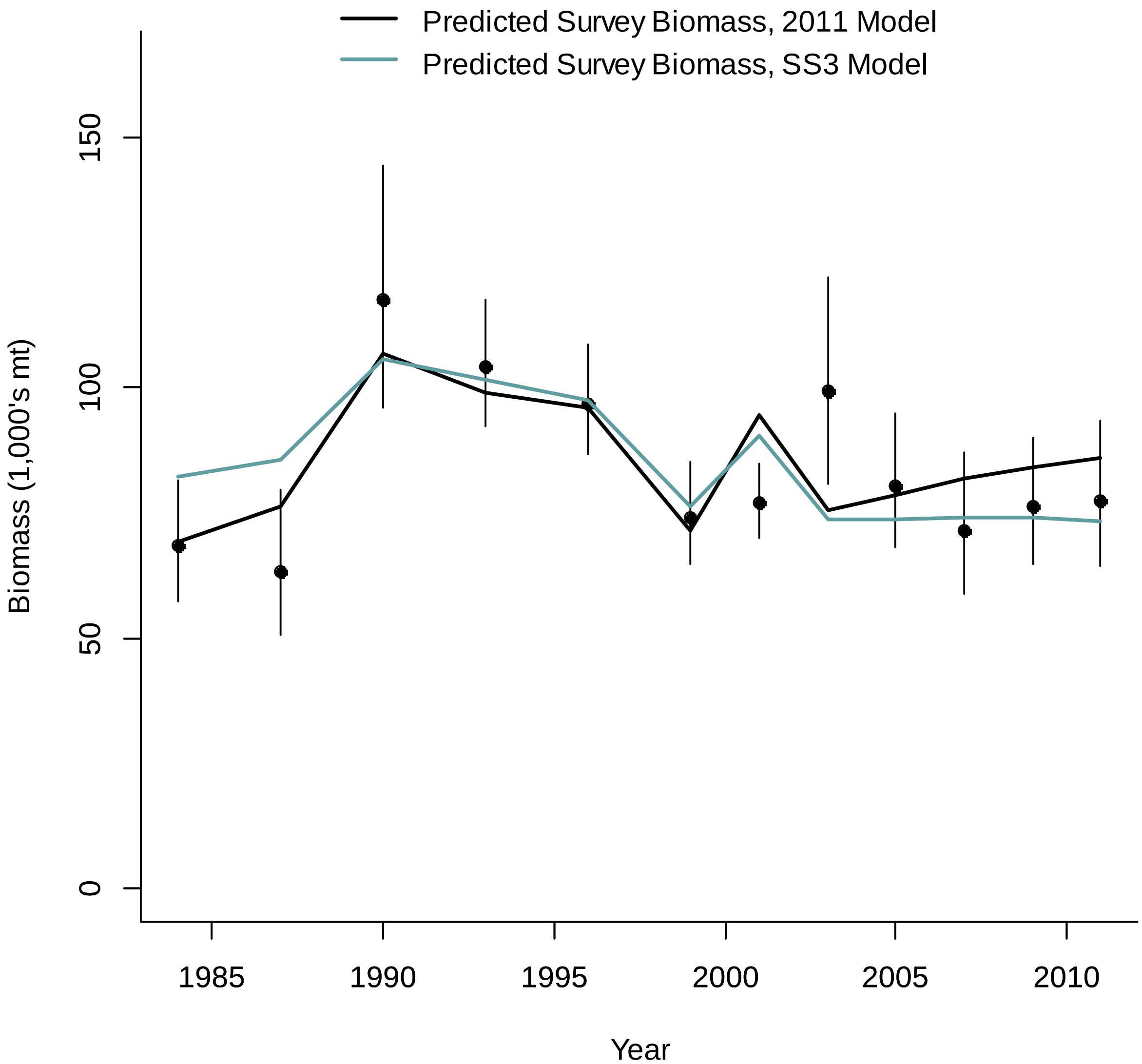
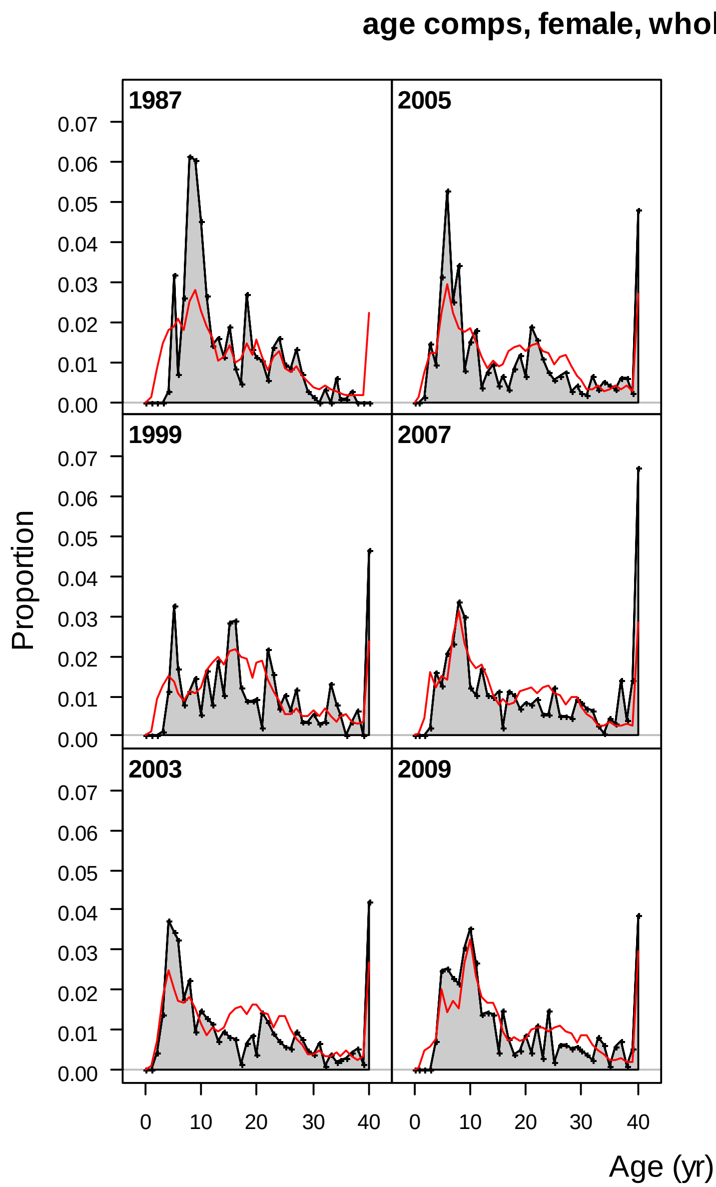


Figure 43. Observed survey biomass (black dots) with 95% asymptotic confidence intervals (vertical black lines) and predicted survey biomass for the proposed alternative model M5 (as for the transitional SS3 model, but with fixed internal growth parameters, estimated initial equilibrium F, and length-based, logistic, sex-specific fishery AND survey selectivity; blue lines) and the 2011 model (black lines).


age comps, female, whole catch, Survey1
1987
2005
1999
2007
2003
2009
Proportion
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0
10
20
30
40
Age (yr)

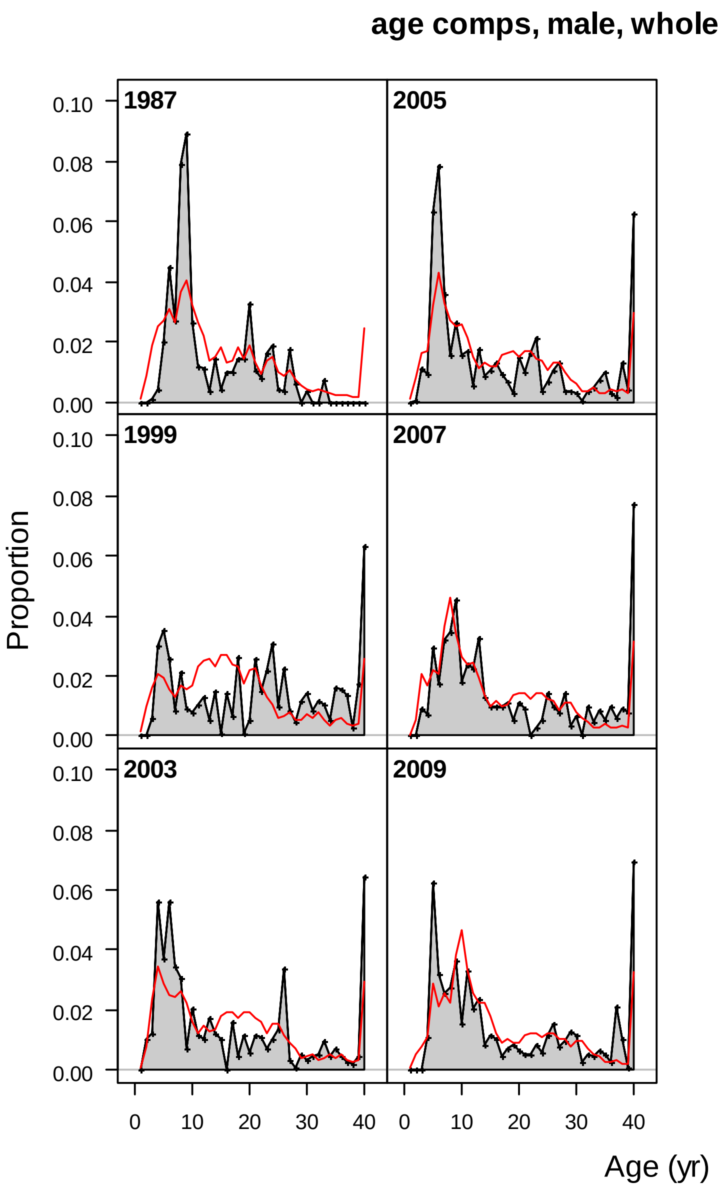
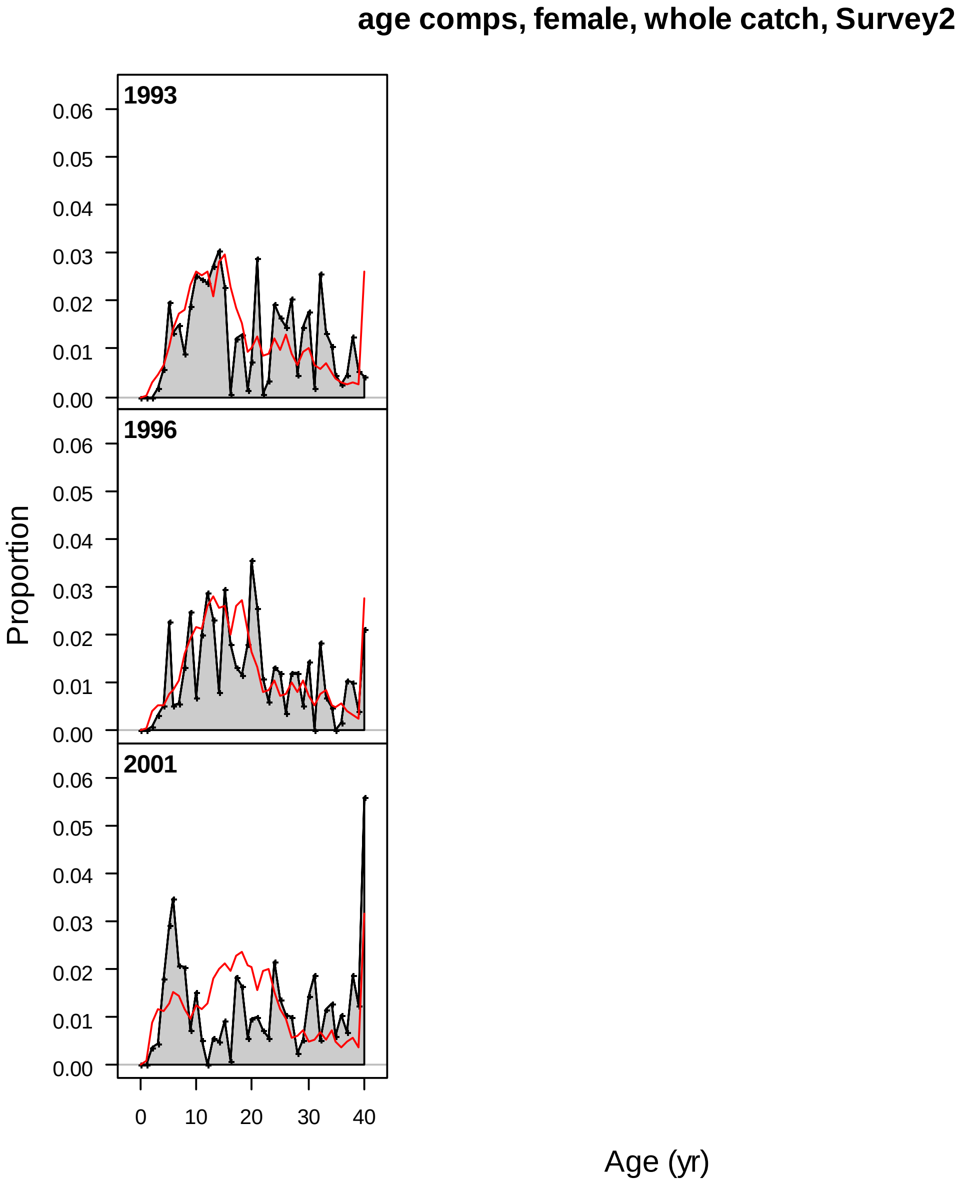


Figure 44. Observed (solid black lines and grey shaded area) and predicted (red lines) full-coverage survey proportions-at-age for proposed alternative model M5 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery AND survey selectivity are logistic, sex-specific, length-based functions) for females (first panel) and males (second panel).


age comps, female, whole catch, Survey2
1993
1996
2001
Proportion
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
10
20
30
40
Age (yr)

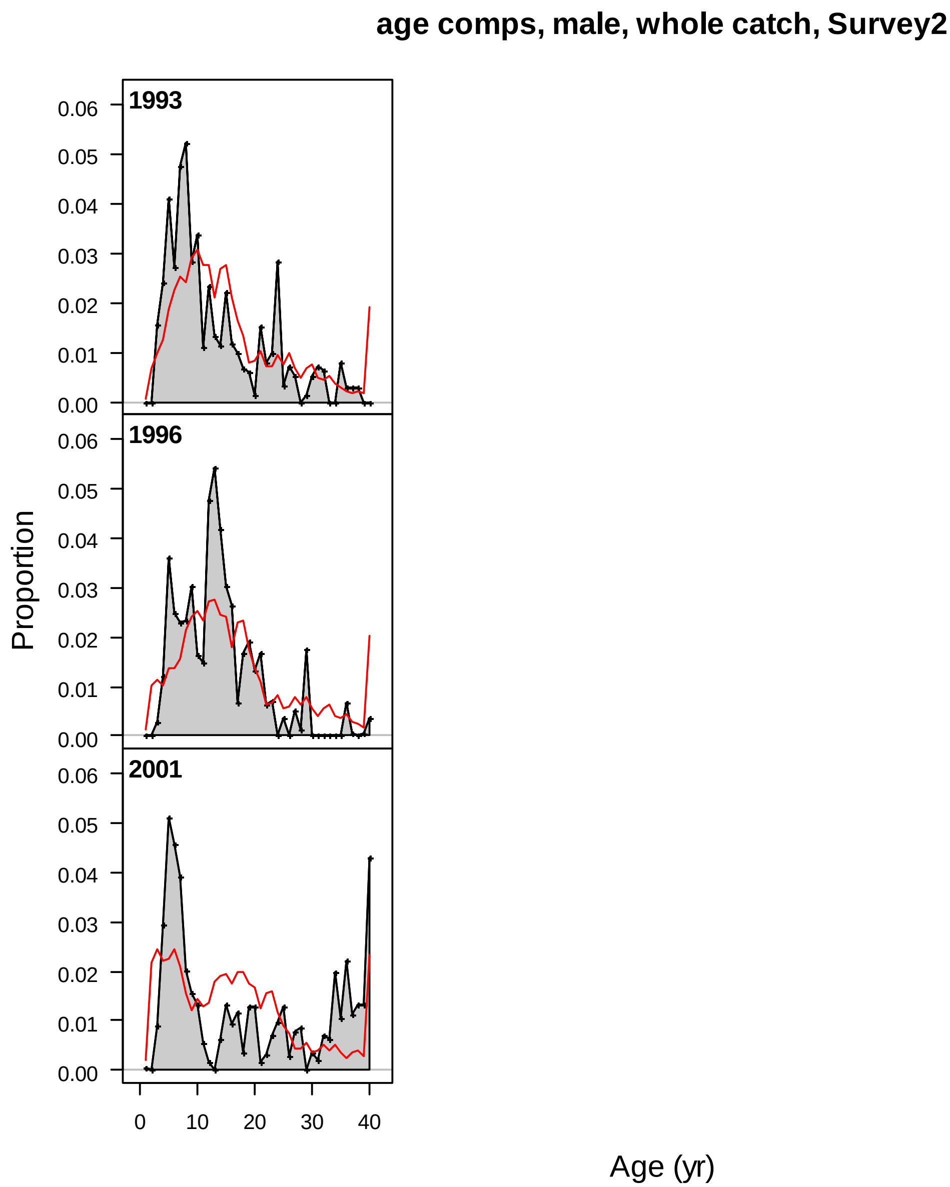
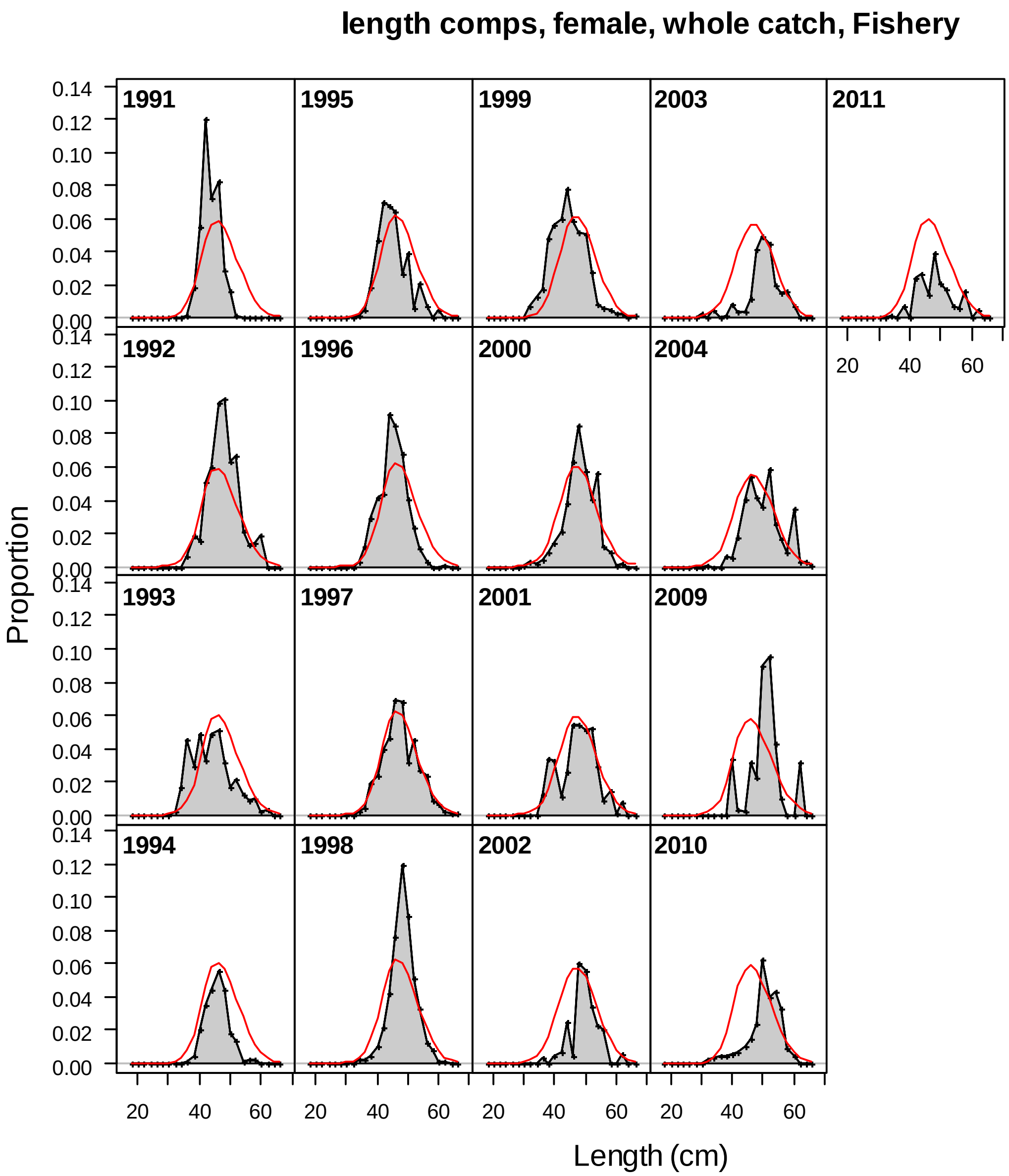


Figure 45. Observed (solid black lines and grey shaded area) and predicted (red lines) shallow-water survey proportions-at-age for proposed alternative model M5 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery AND survey selectivity are logistic, sex-specific, length-based functions) for females (first panel) and males (second panel).


length comps, female, whole catch, Fishery
Proportion
Length (cm)
1991
1995
1999
2003
2011
1992
1996
2000
2004
1993
1997
2001
2009
1994
1998
2002
2010
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
20
40
60

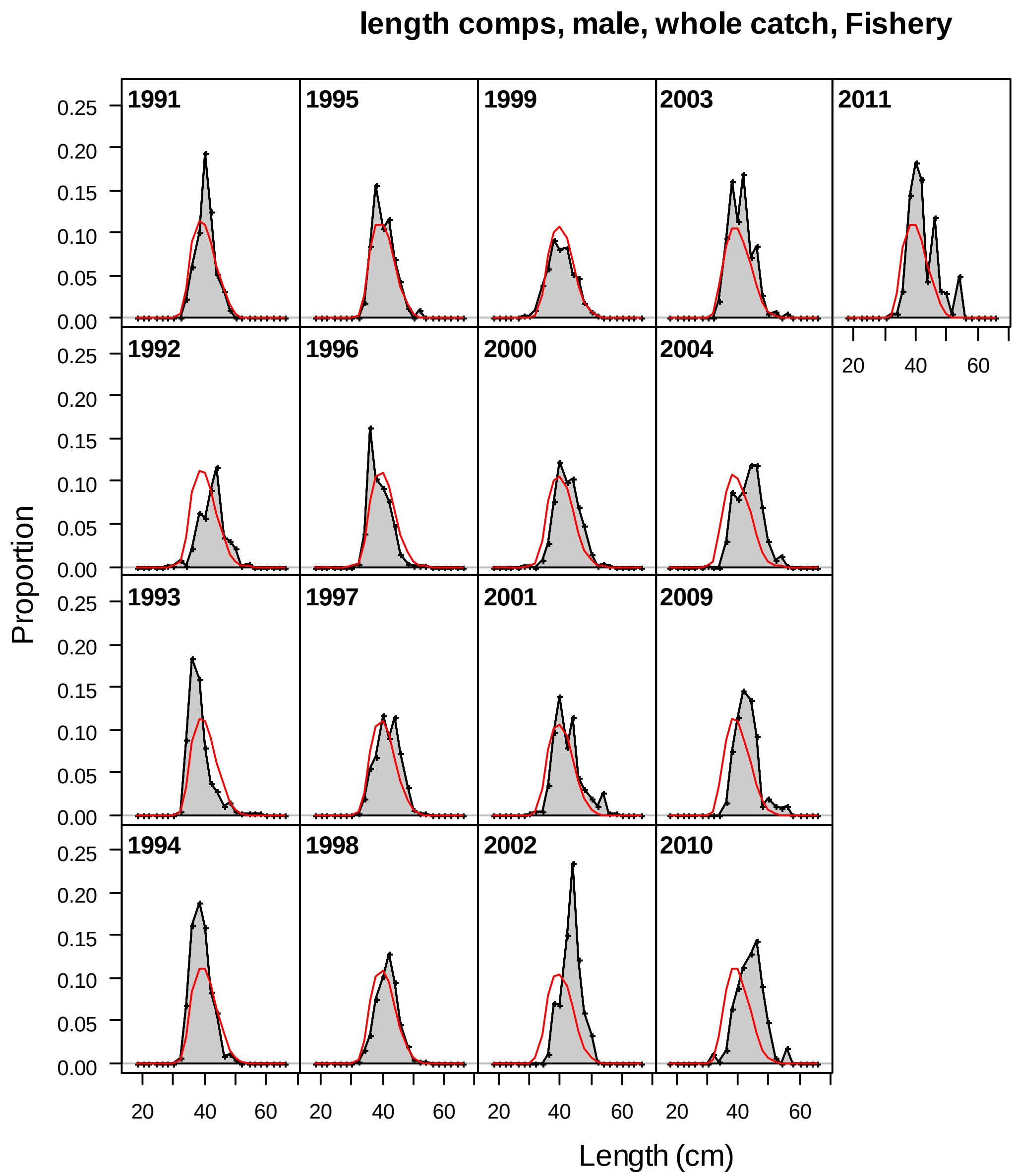
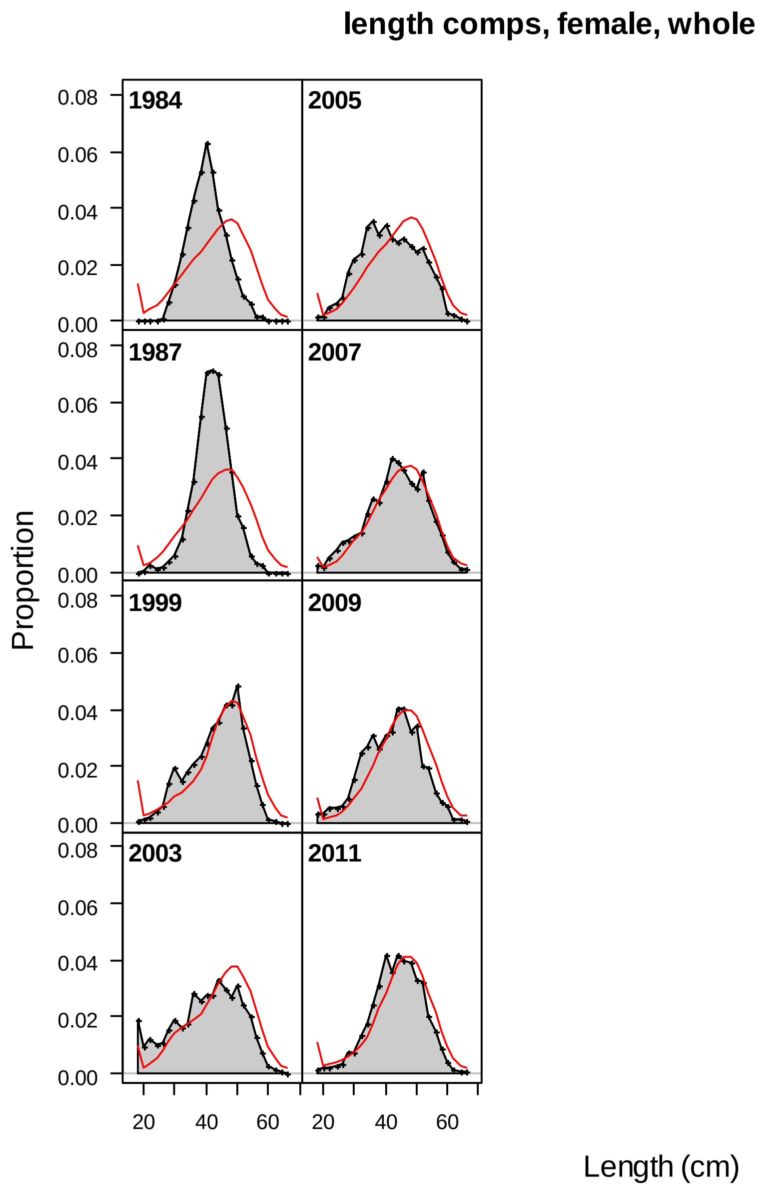


Figure 46. Observed (solid black lines and grey shaded area) and predicted (red lines) fishery proportions-at-length for proposed alternative model M5 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery AND survey selectivity are logistic, sex-specific, length-based functions) for females (first panel) and males (second panel).


length comps, female, whole catch, Survey1
1984
2005
1987
2007
1999
2009
2003
2011
0.00
0.02
0.04
0.06
0.08
Proportion
20
40
60
Length (cm)

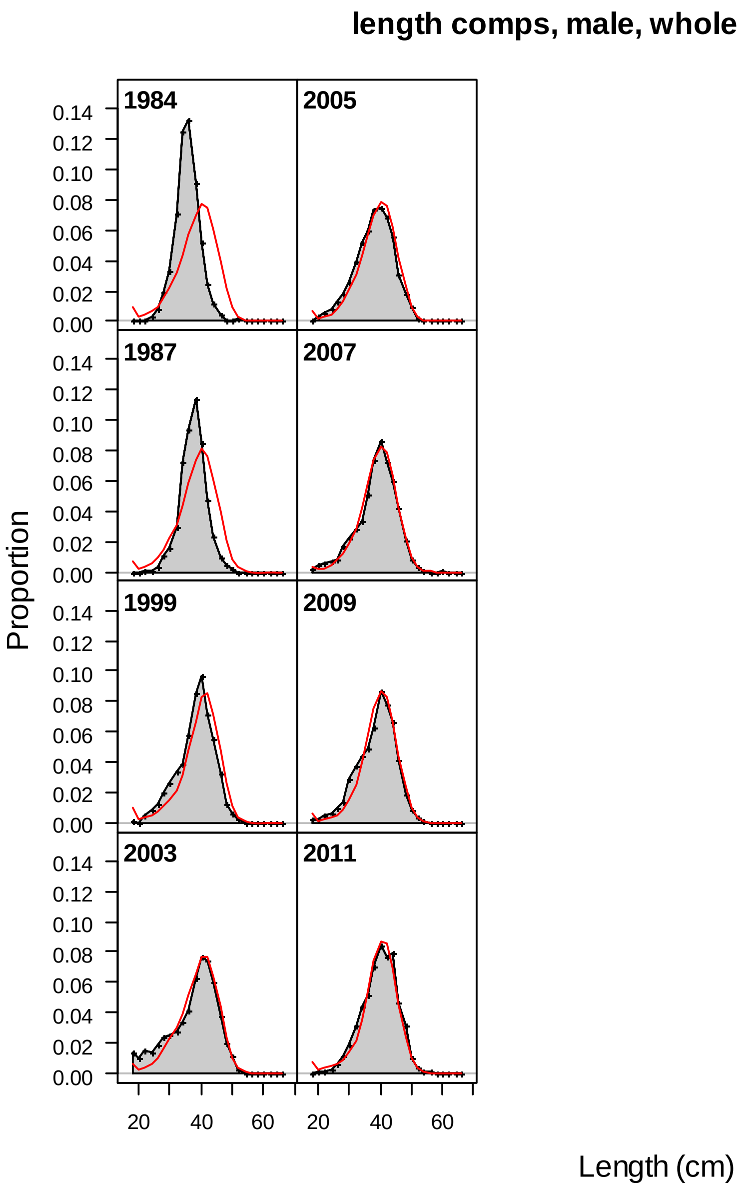
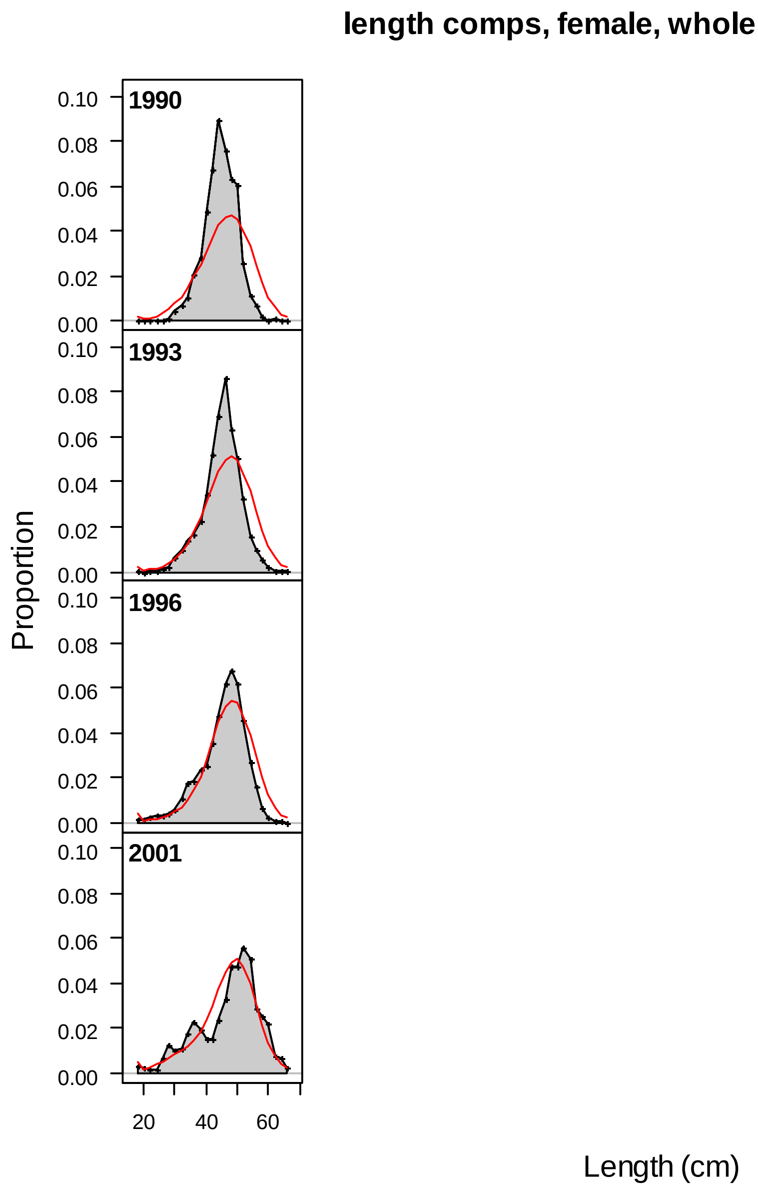


Figure 47. Observed (solid black lines and grey shaded area) and predicted (red lines) full-coverage survey proportions-at-length for proposed alternative model M5 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery AND survey selectivity are logistic, sex-specific, length-based functions) for females (first panel) and males (second panel).


length comps, female, whole catch, Survey2
1990
1993
1996
2001
0.10
0.08
0.06
0.04
0.02
0.00
Proportion
20
40
60
Length (cm)

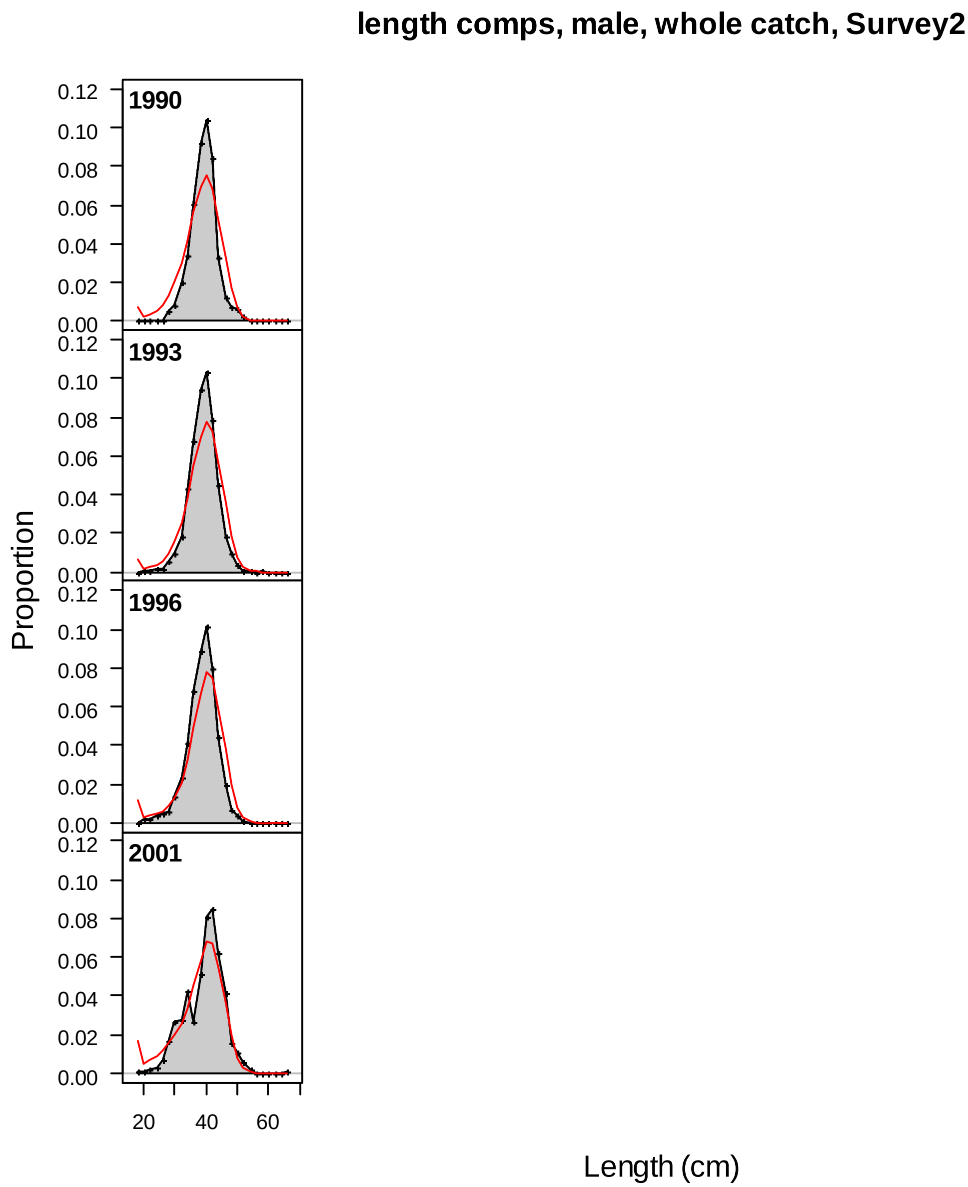


Figure 48. Observed (solid black lines and grey shaded area) and predicted (red lines) shallow-water survey proportions-at-length for proposed alternative model M5 (where growth parameters are specified internally, an initial equilibrium fishing mortality rate is calculated, and fishery AND survey selectivity are logistic, sex-specific, length-based functions) for females (first panel) and males (second panel).